Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze della Terra, della Vita e dell'Ambiente

Ciclo XXX

Settore Concorsuale: 05/B1 - Zoologia e Antropologia

Settore Scientifico Disciplinare: BIO/08 - Antropologia

ARCHAEOLOGICAL GENETICS: A PRELIMINARY OVERVIEW OF THE IRON AGE ITALIAN POPULATION

Presentata da: **Patrizia Serventi**

Coordinatore Dottorato
Prof. Giulio Viola

Supervisore
Prof.ssa Donata Luiselli

Co-Supervisori
Dr.ssa Elisabetta Cilli
Prof.ssa Anna Chiara Fariselli

Esame Finale – Anno 2018

Table of contents

Figures legends

Tables legends

List of abbreviations

A	*Adenine*
AD	*Anno Domini*
aDNA	*ancient DNA*
AMS	*Accelerator Mass Spectrometry*
BC	*Before Christ*
bp	*base pairs*
BSA	*Bovine Serum Albumin*
C	*Cytosine*
CCD	*Charge-Coupled Device*
CO_2	*Carbon dioxide*
CODIS	*Combined DNA Index System*
ddNTP	*DiDeoxyNucleotideTriphosphates*
DI	*Degradation index*
DNA	*DideoxyriboNucleicAcid*
dNTP	*DeoxyNucleotideTriphosphates*
EDTA	*EthyleneDiamineTetraaceticAcid*
ESS	*European Standard Set*
Fst	*Fixation index*
Fwd	*Forward*
g	*gram*
G	*Guanine*
H_2	*Hydrogen*
HCl	*Hydrogen Chloride*
hg	*Haplogroup*
HG	*Hunter-Gatherer*
HTS	*High-Throughput-Sequencing*
HVS	*Hyper Variable Segment*
IBD	*Identical-by-descent*
Indel	*Insertion-Deletion Polymorphisms*
IPC	*Internal PCR Control*
ISP	*Ion Sphere Particle*
kya	*Thousand years ago*
LA	*Large Amplicon*
LBK	*Linear BandKeramik culture*
LD	*Linkage disequilibrium*
LGM	*Last Glacial Maximum*
LR	*Likelihood Ratio*
Mb	*Megabase*
MDS	*Multidimensional Scaling*
mg	*milligram*
$MgCl_2$	*Magnesium chloride*
mL	*millilitre*
mM	*millimolar*
mtDNA	*mitochondrial DNA*
Myr	*Millions of years*
NaClO	*Sodium hypochlorite*
NaOH	*Sodium hydroxide*
ng	*nanogram*
NGS	*Next-Generation Sequencing*
nm	*nanometre*
nM	*nanomolar*
np	*Nucleotide Position*
nuDNA	*nuclear DNA*
PC	*Principal Component*
PCR	*Polymerase Chain Reaction*
pH	*potential of Hydrogen*
pM	*picomolar*
PMI	*post-mortem interval*
qPCR	*Real time PCR*
rCRS	*revised Cambridge Reference Sequence*
Rvs	*Reverse*
SA	*Small Amplicon*
SAP	*Shrimp Alkaline Phosphatase*
SBE	*Single-Base Extension*
SiO_2	*Silicon dioxide*
SNP	*Single Nucleotide Polymorphism*
STR	*Short Tandem Repeat*
T	*Thymine*
TBE	*Tris-Borate-EDTA*
TET	*Tris-EDTA-Tween*
U	*Uracil*
UV	*Ultra Violet*
V	*volt*
ya	*years ago*
yBP	*years before past*
µL	*microlitre*

Sampling step in the ancient DNA laboratory (University of Bologna)

CHAPTER 1

Ancient DNA: an overview

The study of DNA isolated from old biological materials, such as specimens recovered from archaeological finds, museum specimens and fossil record, is a relatively new and rapidly evolving field of research. Early analyses of ancient DNA (aDNA) focused on organellar DNA (mitochondrial in animals and chloroplast in plants) because these are present in multiple copies in the cells and are covered by extra membranes of protection, making preservation, isolation, and analyses much easier. Within the last decade, however, with the advent of more rapid high-throughput DNA sequencing (HTS) technologies, it has become possible to analyse the more informative nuclear genome of a larger number of ancient samples. Today, technological advances allow scientists to read billions of letters from the genomes of ancient humans and other organisms, transforming our view of history and evolution.

1.1 History of ancient DNA study

The first aDNA study was published 25 years ago, when Russell Higuchi and co-workers reported in Nature (Higuchi et al., 1984) the molecular cloning of short mitochondrial DNA (mtDNA) sequence fragments (229 base pairs – bp) from a piece of dry tissue of 140-year-old quagga museum specimen, a zebra-like species (*Equus quagga*) that became extinct in 1883. The quagga work captured international attention, demonstrating how preserved old tissues can retain amplifiable DNA sequences. This finding was shortly followed by a report of the first detection of human nuclear DNA (nuDNA) in an extract of muscle from an infant boy of a pre-Dynastic Egyptian mummy (5th century BC) (Pääbo, 1985a, b,1986). These first approach that yielded aDNA sequences by the cloning of end-repaired DNA molecules into *in*

vivo vectors, such as phages (Higuchi et al., 1984) or bacteria (Pääbo, 1985), showed that the genetic material surviving in ancient specimens was often principally microbial or fungal, and that endogenous DNA was generally limited to very low concentrations of short, damaged fragments of multi-copy loci, such as mtDNA (Pääbo, 1989). In fact, after an organism dies, genetic material is damaged under the effect of cellular nucleases, microbial enzymes and physical factors (Dabney et al., 2013a). As a result, aDNA sequences contain chemical modifications, including strand breaks, DNA crosslinks, and modified bases, that make their recovery challenging (Pääbo, 1989; Lindahl, 1993).

A few years later, the development of the polymerase chain reaction (PCR) technique (Mullis et al., 1986; Mullis and Faloona, 1987), made possible to amplify minute amounts of specific genomic targets up to a level compatible with downstream sequencing (Pääbo, 1989; Pääbo and Wilson 1988; Pääbo et al., 1989; Thomas et al., 1989). Certainly, without the invention of PCR, it is unlikely that ancient DNA research would ever have resulted in more than a few reports of short DNA fragments with little biological significance (Shapiro & Hofreiter, 2014). Neverless, the combination of the high sensitivity of the PCR reaction and the damaged nature of aDNA made modern contamination a crucial problem for these studies. In fact, due to the *postmortem* DNA degradation processes, there is a high chance that an error is introduced in the early cycles of the PCR reaction, thus increasing the risk for preferential amplification of exogenous contaminant sequences (Hofreiter et al., 2001; Rizzi et al., 2012). Proof of this problematic is that many of the first works published have later been proved to be the results of modern contamination, such as early DNA sequences surviving for millions of years (Myr) in plants (Golenberg et al., 1990; Soltis et al., 1992; Kim et al., 2004), dinosaur bones (Woodward et al., 1994) and amber inclusions (Cano et al., 1992a, b, 1993; DeSalle et al., 1992, 1993; Poinar et al., 1993; DeSalle, 1994). Thus, researchers have outlined a series of guidelines to ensure the quality of aDNA data and the reliability of consequent conclusions (Rizzi et al., 2012). Afterwards, these guidelines have gradually evolved into a more detailed and extensive list of requirements, resulting in a rigorous set of criteria for aDNA field (Austin et al., 1997; Cooper and Poinar 2000; Hofreiter et al., 2001; Poinar, 2003; Gilbert et al., 2005; Willerslev and Cooper, 2005; Knapp et al., 2015).

Although the retrieval of multi-copy DNA sequences was often possible, the study of single-copy nuclear DNA from diploid organisms was particularly tough due to high rates of nucleotide damage, short DNA fragment lengths, low endogenous DNA content and the possibility of modern contamination. Within the last decade, with the advent of high-

throughput DNA sequencing (HTS) technologies, many of the issues with endogenous aDNA retrieval have been at least partially overcome. Whereas PCR method allows to amplify a limited number of specific DNA targets at a time, the HTS combines amplification and sequencing of up to several billions of individual DNA library templates at a time (Llamas et al., 2017), thus reducing the sequencing costs. In HTS method, the extracts are used to yield DNA libraries that can then be sequenced or used to isolate DNA fragments of interest by hybridization capture (Burbano et al., 2010). One of the most relevant advantages coming from HST is that short molecules (<50 bp) can be studied. As described above, there is the possibility to enrich the endogenous DNA fraction from highly contaminated aDNA extracts. One of the most popular enrichment approach is the selective capture of regions of interest by hybridisation of aDNA with pre-designed oligonucleotide probes (Burbano et al., 2010; Avila-Arcos et al., 2011; Vilstrup et al., 2013; Carpenter et al., 2013; Fu et al., 2013; Enk et al., 2014; Haak et al., 2015; Paijmans et al., 2016). The targets of such selective capture assays can be complete mitochondrial genomes (Brotherton et al., 2013; Llamas et al., 2016; Posth et al., 2016), genome-wide SNPs (Haak et al., 2015; Mathieson et al., 2015; Fu et al., 2016), exomes (Castellano et al., 2014), chromosomes (Fu et al., 2013), or complete genomes (Carpenter et al., 2013; Enk et al., 2014). These advances allow aDNA researchers to generate a huge amount of data that were inconceivable using previous techniques (Marciniak and Perry, 2017) (Figure 1.1.1; Figure 1.1.2).

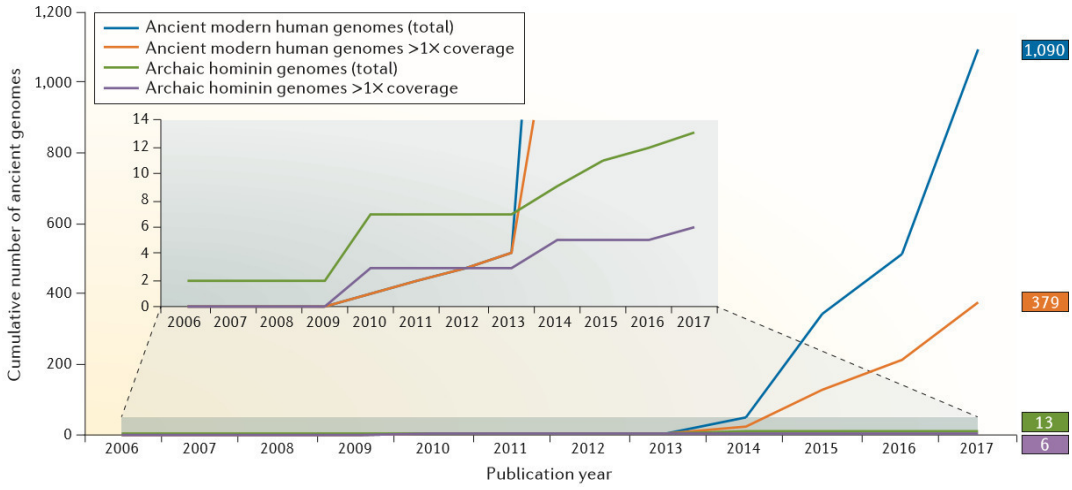


Figure 1.1.1 | Cumulative counts of the numbers of ancient human and archaic hominin individuals with available ancient genomic data (whole genome sequences, exomes, and genome-scale single-nucleotide polymorphism data sets), by year of publication. For each category, both the total number of individuals and the subset of that sample with an average of >1 sequence read per targeted site are depicted (from Marciniak and Perry, 2017).

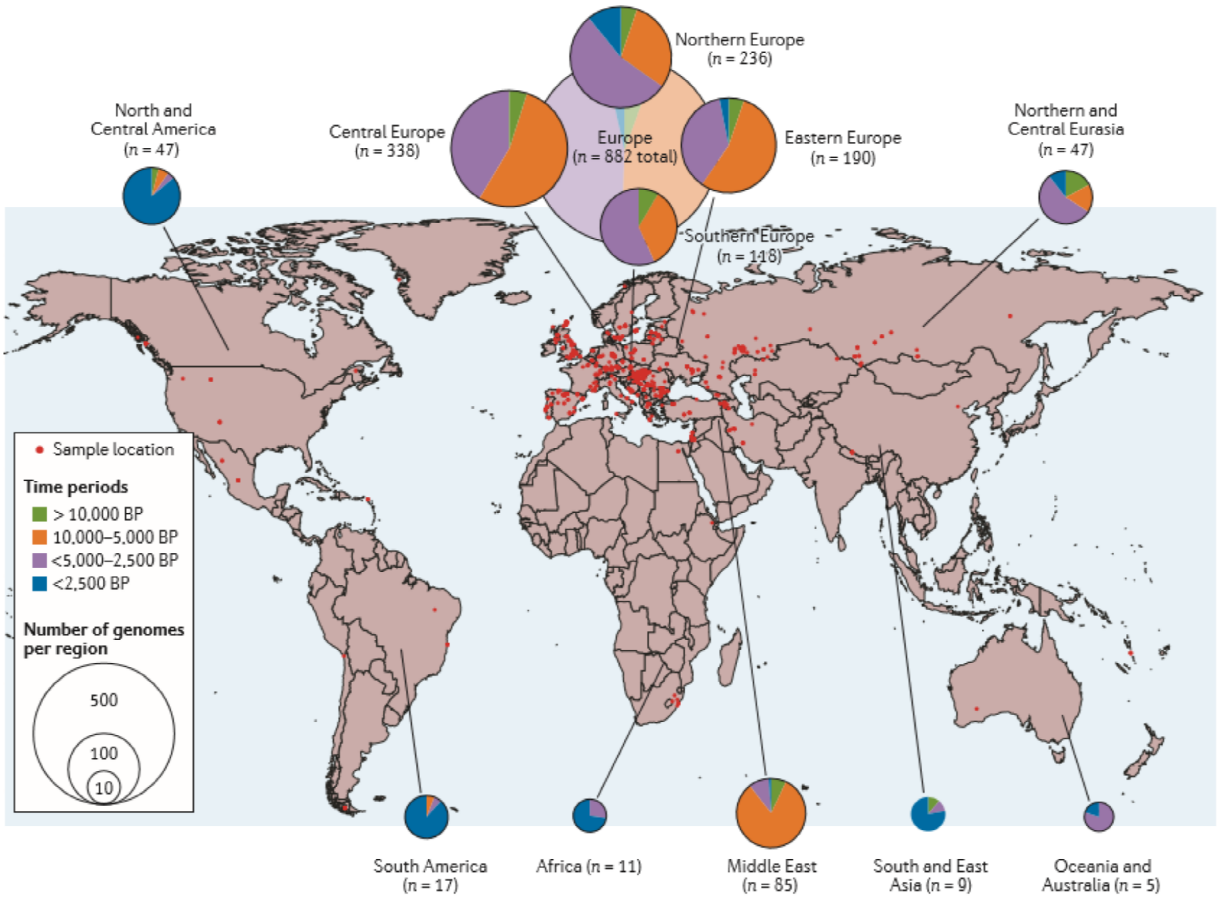


Figure 1.1.2 | A spatiotemporal distribution of ancient human and archaic hominin genome data sets, where the areas of the circles are proportional to the genome counts. For some archaeological sites, precise geographic coordinates were not indicated in the corresponding publication; longitude and latitude were estimated in these cases (from Marciniak and Perry, 2017).

Refinements of the techniques that allow short DNA sequences to be extracted efficiently (Rohland and Hofreiter, 2007; Dabney et al., 2013), as well as the finding that DNA is particularly likely to survive in the petrous part of the temporal bones of humans and animals (Gamba et al., 2014; Pinhasi et al., 2015), and also teeth (Higgins et al., 2013; Gamba et al., 2014, Damgaard et al., 2015; Hansen et al., 2017), have made it possible to retrieve genome-wide DNA data from large numbers of remains.

1.2 The aDNA challenges

1.2.1 Degradation of ancient biomolecules

An understanding of the degradation of ancient molecules is an essential complement to the use of these molecules to address archaeological question. Most recoverable fragments of aDNA are shorter than 150 bp and contain miscoding lesions that can result in erroneous

sequences (Prüfer et al., 2010; Sawyer et al., 2012). The nucleic acids *postmortem* instability is caused by a wide range of degradation reactions that result in the fragmentation and chemical modification of DNA templates (Figure 1.2.1.1).

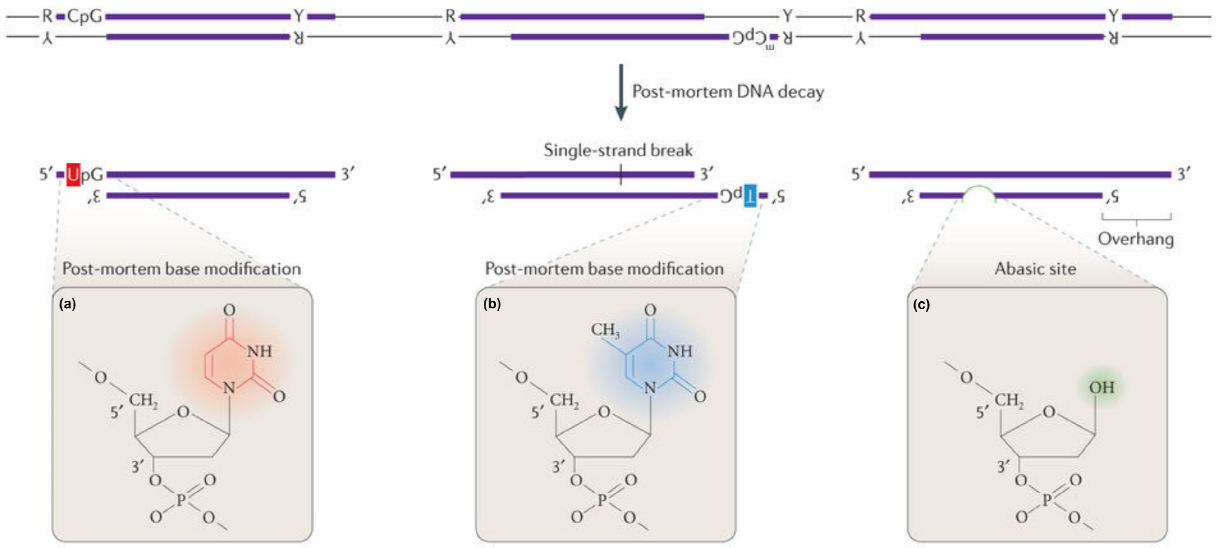


Figure 1.2.1.1 | Typical ancient DNA molecules. The most common base modification identified is deamination of cytosine into uracils (a), or thymines (b) when cytosines were methylated (mC). Such deaminations occur much faster at break ends. Other modification include abasic site (c) (from Orlando et al., 2015).

After the death of an organism, DNA repair mechanisms are blocked. During the life of the host, these mechanisms ensure an effective replication of DNA, but without them, the DNA naturally and gradually degrades into short fragments. Initially, endogenous nucleases start the process of DNA degradation, after that a combination of exogenous nucleases (released by microorganisms and environmental invertebrates) and environmental processes, such as exposure to oxygen and water, carry on the damage process.

Compared to the previously described enzymatic degradation processes, non-enzymatic reactions are slower but persistent. The depurination, which is a hydrolytic damage and the principal mechanism of DNA degradation, falls into this category. It consists of the hydrolysis of an N-glycosidic bounds base-sugar resulting in a base loss. In correspondence of this abasic site, the strand breaks more easily. In details, purines are liberated from DNA at similar rates with guanine being released slightly more rapidly, whereas pyrimidines are lost at 5% of the rate of the purines. At a site of base loss, the DNA chain is weakened and undergoes strand cleavage by β-elimination (Lindhal, 1993).

The best characterized of degradation reaction is the deamination, another hydrolytic damage, that also induce miscoding lesions that cause base misincorporations. The most

common deamination reaction is from cytosine to uracil, thus resulting in a C → T transition. Deamination of purines, even if it occurs, is a minor reaction.

Hydroxyl radicals cause oxidative lesions that generate modifications in sugars and in bases, resulting in base misincorporations and/or block of DNA replication (Dabney et al., 2013b). In detail, the more common mutagenic lesions are the formation of an 8-hydroxyguanine, which base-pairs with adenine rather than cytosine, and of a ring saturated derivative of a pyrimidine, which occurs in several forms and results in noncoding bases. Moreover, abasic sites may proceed to DNA crosslinks between DNA and proteins or between the ring-opened sugar of the abasic sugar and an amino group located on the opposite strand and may prevent the amplification of endogenous DNA. A variety of lesions including oxidative damages, single and double-strand breaks, base modifications, destruction of sugars, intra and interstrand crosslinks and formation of dimers could be caused by radiation. All these decay processes are responsible for the characteristics of aDNA recovered from the samples and they begin almost immediately after the organism dies.

Researchers have assumed that DNA survival correlates negatively with the age of the material. Nevertheless, several studies suggest that more than age, it is the sedimentary environment that influences the rate of decay of a sample (Pääbo et al., 1989; Höss et al., 1996; Poinar et al., 1996). Constant low temperatures, rapid desiccation, the absence of microorganisms, high salt concentrations, oxygen absence, neutral or slightly alkaline environments have the possibility to enhance preservation of ancient DNA. In addition, few environmental situations can properly exclude the microbial activity, for example, if DNA binds to mineral surfaces and even fermentation activities of anaerobic bacteria and presence of humic acids (Daskalaki, 2014). Bollongino et al. (2008) also suggested that fresh excavated samples show a better preservation state (evaluated as PCR successful rate) than samples stored in the museum. As a result, extracted aDNA is always a mixture of endogenous and exogenous DNA, including DNA from bacteria, fungi, and other organisms that colonize the sample during burial, and any contamination occurring during excavations and processing from people who handle the specimen (Shapiro and Hofreiter, 2014).

1.2.2 aDNA Contamination

Given the conditions described above, the endogenous DNA is often contaminated with some level of exogenous DNA, such as bacteria, fungi, viruses, algae, *postmortem* juxtaposition of organisms, or modern human DNA from the researchers themselves

(Morozova et al., 2016). PCR-based studies had shown the extent of human contamination introduced during handling of bone and tooth samples when stringent aDNA precautions are not in place (Gilbert et al., 2005; Sampietro et al., 2006; Pilli et al., 2013). For example, laboratory instruments and reagents could be contaminated by modern human DNA during their production, before arriving into an aDNA laboratory (Leonard et al., 2007; Champlot et al., 2010; Deguilloux et al., 2011).

In order to avoid contamination, the experiment must be properly managed, including special requirements for sample collection, sterilization of the working area, DNA authentication, and independent reproducibility (Cooper and Poinar, 2000; Knapp et al., 2015; Llamas et al., 2017). These protocols are constantly being refined and improved. For example, there are several guidelines for sampling (Brown and Brown, 1992; Yang and Watt, 2005; Pruvost et al., 2007; Fortea et al., 2008; Pilli et al., 2013; Llamas et al., 2017), which are equally relevant for archaeologists working in the field, physical anthropologists and museum curators who handle the remains once unearthed. Regarding the protocols used for working in the aDNA laboratory (Cooper and Poinar, 2000; Poinar, 2003; Pääbo et al., 2004; Champlot et al., 2010; Knapp et al., 2012), in addition to mechanical removal of the outer layer and UV and/or Sodium hypochlorite treatment of the sample, a brief pre-digestion step was recently suggested, consisting of short-term sample incubation (15-30 minutes) in an extraction buffer and its subsequent removal (Kemp and Smith, 2005; Salamon et al., 2005; Malmstrom et al., 2007; Korlević et al., 2015; Damgaard et al., 2015). According to the authors, this step alone increases the fraction of endogenous DNA several fold (Damgaard et al., 2015; Allentoft et al., 2015). Irrespective of which decontamination method is used, aDNA researchers have to balance the removal of contaminating DNA with preserving the remaining endogenous DNA (Llamas et al., 2017).

1.3 The applications of biomolecular archaeology

Following the introduction of these technical improvements, the expertise in aDNA field has expanded at a rapid pace. aDNA allows to observe changes in genetic diversity through time and and geography and therefore can be used to test a hypothesis about the relationships between environmental events and evolutionary changes in populations, to resolve controversy about evolutionary relationships between species, to reveal otherwise cryptic relationships between past and present population. At a genomic level, aDNA studies can identify regions under selection within a genome, including genetic changes that may underlie

species-specific traits and provide a tool to investigate genome evolution, including the evolution of pathogens response (Bos et al., 2011; 2015; Valtuena et al., 2017). Moreover, paleogenomes whose ages are well constrained may be useful to calibrate a molecular clock or to investigate genome stability (Shapiro and Hofreiter, 2014).

For archaeological studies, aDNA has the opportunity to resolve questions that classical anthropology could just partially explain. Human population history and evolution is a broad area where valuable studies have been carried out. Thanks to the HTS method it was possible to sequence the complete mitochondrial genome, hundreds of thousands of nuclear variants, and the nuclear genome from several hundreds of archaeological human remains (Llamas et al., 2017), for example, the genome sequences of Neanderthals (Green et al., 2010; Prüfer et al., 2013, 2017), the archaic hominins from Sima de los Huesos in Spain (Meyer et al., 2013, 2016), or from the Denisova Cave in Russia (Meyer et al., 2012).

In the latest aDNA researches the attention has been focused on population genetic studies (Raghavan et al., 2013; Allentoft et al., 2015; Mathieson et al., 2015; Haak et al., 2015; Fu et al., 2016; Lazaridis et al., 2016, 2017; Haber et al., 2017; Margaryan et al., 2017; Schuenemann et al., 2017), as well as determed genetic relationship of samples (Vohr et al., 2015), genetic sex determination (Green et al., 2010; Skoglund et al., 2013; Sikora et al., 2017), or define kinship, demography, health, subsistence practices, and the social organization of past populations (Haak et al., 2008; Raff et al., 2011; Kirsanow and Burger ,2012; Brandt et al., 2015).

References:

Allentoft ME, Sikora M, Sjögren K-G, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, et al. 2015. Population genomics of Bronze Age Eurasia. Nature 522:167–172.

Andrades Valtueña A, Mittnik A, Key FM, Haak W, Allmäe R, Belinskij A, Daubaras M, Feldman M, Jankauskas R, Janković I, et al. 2017. The Stone Age Plague and Its Persistence in Eurasia. Curr. Biol. 27:3683–3691.

Austin JJ, Smith AB, Thomas RH. 1997. Palaeontology in a molecular world: the search for authentic ancient DNA. Trends Ecol. Evol. 12:303–6.

Ávila-Arcos MC, Cappellini E, Romero-Navarro JA, Wales N, Moreno-Mayar JV, Rasmussen M, Fordyce SL, Montiel R, Vielle-Calzada J-P, Willerslev E, et al. 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Sci. Rep. 1:74.

Bollongino R, Tresset A, Vigne J-D. 2008. Environment and excavation: Pre-lab impacts on ancient DNA analyses. Comptes Rendus Palevol 7:91–98.

Bos KI, Jäger G, Schuenemann VJ, Vågene ÅJ, Spyrou MA, Herbig A, Nieselt K, Krause J. 2015. Parallel detection of ancient pathogens via array-based DNA capture. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 370:20130375.

Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S, et al. 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478:506–510.

Brandt G, Szécsényi-Nagy A, Roth C, Alt KW, Haak W. 2015. Human paleogenetics of Europe – The known knowns and the known unknowns. J. Hum. Evol. 79:73–92.

Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, Richards SM, Sarkissian C Der, Ganslmeier R, Friederich S, et al. 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat. Commun. 4:1764.

Brown TA, Brown KA. 1992. Ancient DNA and the archaeologist. Antiquity 66:10–23.

Burbano HA, Hodges E, Green RE, Briggs AW, Krause J, Meyer M, Good JM, Maricic T, Johnson PLF, Xuan Z, et al. 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328:723–5.

Cano RJ, Poinar HN, Poinar GO, 1992a. Isolation and partial characterisation of DNA from the bee Proplebeia dominicana (Apidae: Hymenoptera) in 25-40 million year old amber. Med. Sci. Res. 20:249–251.

Cano RJ, Poinar HN, Roublik DW, Poinar GO, 1992b. Enzymatic amplification and nucleotide sequencing of portions of the 18S rRNA gene of the bee Proplebeia dominicana (Apidae: Hymenoptera) isolated from 25-40 million year old Dominican amber. Med. Sci. Res. 20:619–622.

Cano RJ, Poinar HN, Pieniazek NJ, Acra A, Poinar GO 1993. Amplification and sequencing of DNA from a 120–135-million-year-old weevil. Nature 363:536–538.

Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, Sikora M, Rasmussen M, Gravel S, Guillén S, Nekhrizov G, et al. 2013. Pulling out the 1%: Whole-Genome Capture for the Targeted Enrichment of Ancient DNA Sequencing Libraries. Am. J. Hum. Genet. 93:852–864.

Castellano S, Parra G, Sánchez-Quinto FA, Racimo F, Kuhlwilm M, Kircher M, Sawyer S, Fu Q, Heinze A, Nickel B, et al. 2014. Patterns of coding variation in the complete exomes of three Neandertals. Proc. Natl. Acad. Sci. U. S. A. 111:6666–71.

Champlot S, Berthelot C, Pruvost M, Bennett EA, Grange T, Geigl E-M. 2010. An Efficient Multistrategy DNA Decontamination Procedure of PCR Reagents for Hypersensitive PCR Applications. Lalueza-Fox C, editor. PLoS One 5:e13042.

Cooper A, Poinar HN. 2000. Ancient DNA: do it right or not at all. Science 289:1139.

Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga J-L, et al. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. 110:15758–15763.

Dabney J, Meyer M, Pääbo S. 2013. Ancient DNA damage. Cold Spring Harb. Perspect. Biol. 5.

Damgaard PB, Margaryan A, Schroeder H, Orlando L, Willerslev E, Allentoft ME. 2015. Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. 5:11184.

Deguilloux M-F, Ricaud S, Leahy R, Pemonge M-H. 2011. Analysis of ancient human DNA and primer contamination: One step backward one step forward. Forensic Sci. Int. 210:102–109.

DeSalle R. 1994. Implications of ancient DNA for phylogenetic studies. Experientia 50:543–50.

DeSalle R, Barcia M, Wray C. 1993. PCR jumping in clones of 30-million-year-old DNA fragments from amber preserved termites (Mastotermes electrodominicus). Experientia 49:906–9.

DeSalle R, Gatesy J, Wheeler W, Grimaldi D. 1992. DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. Science 257:1933–6.

Enk JM, Devault AM, Kuch M, Murgha YE, Rouillard J-M, Poinar HN. 2014. Ancient Whole Genome Enrichment Using Baits Built from Modern DNA. Mol. Biol. Evol. 31:1292–1294.

Fortea J, de la Rasilla M, García-Tabernero A, Gigli E, Rosas A, Lalueza-Fox C. 2008. Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain). J. Hum. Evol. 55:353–357.

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Pääbo S. 2013. DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl. Acad. Sci. U. S. A. 110:2223–7.

Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, et al. 2016. The genetic history of Ice Age Europe. Nature 534:200.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. 5:5257.

Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I. 2005. Assessing ancient DNA studies. Trends Ecol. Evol. 20:541–4.

Gilbert MTP, Willerslev E, Hansen AJ, Barnes I, Rudbeck L, Lynnerup N, Cooper A. 2003. Distribution patterns of postmortem damage in human mitochondrial DNA. Am. J. Hum. Genet. 72:32–47.

Golenberg EM, Giannasi DE, Clegg MT, Smiley CJ, Durbin M, Henderson D, Zurawski G. 1990. Chloroplast DNA sequence from a Miocene Magnolia species. Nature 344:656–658.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-Y, et al. 2010. A draft sequence of the Neandertal genome. Science 328:710–722.

Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, Hawkesworth C, Pike AWG, Meller H, Alt KW. 2008. Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc. Natl. Acad. Sci. U. S. A. 105:18226–31.

Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. Nature 522:207–211.

Hansen HB, Damgaard PB, Margaryan A, Stenderup J, Lynnerup N, Willerslev E, Allentoft ME. 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. Caramelli D, editor. PLoS One 12:e0170940.

Higgins D, Austin JJ. 2013. Teeth as a source of DNA for forensic identification of human remains: A Review. Sci. Justice 53:433–441.

Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. Nature 312:282–284.

Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S. 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 29:4793–9.

Höss M, Jaruga P, Zastawny TH, Dizdaroglu M, Pääbo S. 1996. DNA damage and DNA sequence retrieval from ancient tissues. Nucleic Acids Res. 24:1304–7.

Kemp BM, Smith DG. 2005. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. Forensic Sci. Int. 154:53–61.

Kim S, Soltis DE, Soltis PS, Suh Y. 2004. DNA sequences from Miocene fossils: an ndhF sequence of Magnolia latahensis (Magnoliaceae) and an rbcL sequence of Persea pseudocarolinensis (Lauraceae). Am. J. Bot. 91:615–20.

Kirsanow K, Burger J. 2012. Ancient human DNA. Ann. Anat. 194:121–32.

Knapp M, Lalueza-Fox C, Hofreiter M. 2015. Re-inventing ancient human DNA. Investig. Genet. 6:4.

Korlević P, Gerber T, Gansauge M-T, Hajdinjak M, Nagel S, Aximu-Petri A, Meyer M. 2015. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques 59.

Lazaridis I, Mittnik A, Patterson N, Mallick S, Rohland N, Pfrengle S, Furtwängler A, Peltzer A, Posth C, Vasilakis A, et al. 2017. Genetic origins of the Minoans and Mycenaeans. Nature 548:214.

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. Nature 536:419–424.

Leonard JA, Shanks O, Hofreiter M, Kreuz E, Hodges L, Ream W, Wayne RK, Fleischer RC. Animal DNA in PCR reagents plagues ancient DNA research.

Lindahl T. 1993. Instability and decay of the primary structure of DNA. Nature 362:709–715.

Llamas B, Valverde G, Fehren-Schmitz L, Weyrich LS, Cooper A, Haak W. 2017. From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. STAR Sci. Technol. Archaeol. Res. 3:1–14.

Llamas B, Willerslev E, Orlando L. 2016. Human evolution: a tale from ancient genomes. Philos. Trans. R. Soc. London B Biol. Sci. 372.

Malmstrom H, Svensson EM, Gilbert MTP, Willerslev E, Gotherstrom A, Holmlund G. 2007. More on Contamination: The Use of Asymmetric Molecular Behavior to Identify Authentic Ancient Human DNA. Mol. Biol. Evol. 24:998–1004.

Marciniak S, Perry GH. 2017. Harnessing ancient genomes to study the history of human adaptation. Nat. Rev. Genet. 18:659–674.

Margaryan A, Derenko M, Hovhannisyan H, Malyarchuk B, Heller R, Khachatryan Z, Avetisyan P, Badalyan R, Bobokhyan A, Melikyan V, et al. 2017. Eight Millennia of Matrilineal Genetic Continuity in the South Caucasus. Curr. Biol. 337:957–960.

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528:499–503.

Meyer M, Arsuaga J-L, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, Martínez I, Gracia A, de Castro JMB, Carbonell E, et al. 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. Nature 531:504–507.

Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, Arsuaga J-L, Martínez I, Gracia A, de Castro JMB, Carbonell E, et al. 2013. A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature 505:403–406.

Meyer M, Kircher M, Gansauge M-T, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prüfer K, de Filippo C, et al. 2012. A high-coverage genome sequence from an archaic Denisovan individual. Science 338:222–6.

Morozova I, Flegontov P, Mikheyev AS, Bruskin S, Asgharian H, Ponomarenko P, Klyuchnikov V, ArunKumar G, Prokhortchouk E, Gankin Y, et al. 2016. Toward high-resolution population genomics using archaeological samples. DNA Res. 23:295–310.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 51 Pt 1:263–73.

Mullis KB, Faloona FA. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. 155:335–50.

Olalde I, Brace S, Allentoft ME, Armit I, Kristiansen K, Rohland N, Mallick S, Booth T, Szécsényi-Nagy A, Mittnik A, et al. 2017. The Beaker Phenomenon And The Genomic Transformation Of Northwest Europe. bioRxiv.

Orlando L, Gilbert MTP, Willerslev E. 2015. Reconstructing ancient genomes and epigenomes. Nat. Rev. Genet. 16:395–408.

Pääbo S. 1986. Molecular genetic investigations of ancient human remains. Cold Spring Harb. Symp. Quant. Biol. 51 Pt 1:441–6.

Pääbo S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc. Natl. Acad. Sci. U. S. A. 86:1939–43.

Pääbo S. Molecular cloning of Ancient Egyptian mummy DNA. Nature 314:644–5.

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. 2004. Genetic Analyses from Ancient DNA. Annu. Rev. Genet. 38:645–679.

Pääbo S, Wilson AC. 1988. Polymerase chain reaction reveals cloning artefacts. Nature 334:387–388.

Paijmans JLA, Fickel J, Courtiol A, Hofreiter M, Förster DW. 2016. Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. Mol. Ecol. Resour. 16:42–55.

Pilli E, Modi A, Serpico C, Achilli A, Lancioni H, Lippi B, Bertoldi F, Gelichi S, Lari M, Caramelli D. 2013. Monitoring DNA Contamination in Handled vs. Directly Excavated Ancient Human Skeletal Remains. Lalueza-Fox C, editor. PLoS One 8:e52524.

Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, Gerritsen F, Moiseyev V, Gromov A, Raczky P, et al. 2015. Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLoS One 10:e0129102.

Poinar HN, Poinar GO, Cano R.J, 1993. DNA from an extinct plant. Nature. 1993;363:677.

Poinar H, Kuch M, McDonald G, Martin P, Pääbo S. 2003. Nuclear gene sequences from a late pleistocene sloth coprolite. Curr. Biol. 13:1150–2.

Poinar HN, Höss M, Bada JL, Pääbo S. 1996. Amino acid racemization and the preservation of ancient DNA. Science 272:864–6.

Posth C, Renaud G, Mittnik A, Drucker DG, Rougier H, Cupillard C, Valentin F, Thevenet C, Furtwängler A, Wißing C, et al. 2016. Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. Curr. Biol.

Prüfer K, de Filippo C, Grote S, Mafessoni F, Korlević P, Hajdinjak M, Vernot B, Skov L, Hsieh P, Peyrégne S, et al. 2017. A high-coverage Neandertal genome from Vindija Cave in Croatia. Science 358:655–658.

Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C, et al. 2013. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature 505:43–49.

Prüfer K, Stenzel U, Hofreiter M, Pääbo S, Kelso J, Green RE. 2010. Computational challenges in the analysis of ancient DNA. Genome Biol. 11:R47.

Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, Morel N, Fernandez-Jalvo Y, Grange T, Geigl E-M. 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. Proc. Natl. Acad. Sci. U. S. A. 104:739–44.

Raff JA, Bolnick DA, Tackney J, O'Rourke DH. 2011. Ancient DNA perspectives on American colonization and population history. Am. J. Phys. Anthropol. 146:503–514.

Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford Jr TW, Orlando L, Metspalu E, et al. 2013. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature 505:87–91.

Rizzi E, Lari M, Gigli E, De Bellis G, Caramelli D. 2012. Ancient DNA studies: new perspectives on old samples. Genet. Sel. Evol. 44:21.

Rohland N, Hofreiter M. 2007. Comparison and optimization of ancient DNA extraction. Biotechniques 42:343–52.

Salamon M, Tuross N, Arensburg B, Weiner S. 2005. Relatively well preserved DNA is present in the crystal aggregates of fossil bones. Proc. Natl. Acad. Sci. U. S. A. 102:13783–8.

Sampietro ML, Gilbert MTP, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006. Tracking down Human Contamination in Ancient Human Teeth. Mol. Biol. Evol. 23:1801–1807.

Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012. Temporal Patterns of Nucleotide Misincorporations and DNA Fragmentation in Ancient DNA. Lalueza-Fox C, editor. PLoS One 7:e34131.

Schuenemann VJ, Peltzer A, Welte B, van Pelt WP, Molak M, Wang C-C, Furtwängler A, Urban C, Reiter E, Nieselt K, et al. 2017. Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat. Commun. 8:15694.

Shapiro B, Hofreiter M. 2014. A Paleogenomic Perspective on Evolution and Gene Function: New Insights from Ancient DNA. Science (80-.). 343:1236573–1236573.

Sikora M, Seguin-Orlando A, Sousa VC, Albrechtsen A, Korneliussen T, Ko A, Rasmussen S, Dupanloup I, Nigst PR, Bosch MD, et al. 2017. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. Science (80-.). 358:659–662.

Skoglund P, Storå J, Götherström A, Jakobsson M. 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. 40:4477–4482.

Soltis PS, Soltis DE, Smiley CJ. 1992. An rbcL sequence from a Miocene Taxodium (bald cypress). Proc. Natl. Acad. Sci. U. S. A. 89:449–51.

Szécsényi-Nagy A, Brandt G, Haak W, Keerl V, Jakucs J, Möller-Rieker S, Köhler K, Mende BG, Oross K, Marton T, et al. 2015. Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. Proc. Biol. Sci. 282:20150339-.

Thomas RH, Schaffner W, Wilson AC, Pääbo S. 1989. DNA phylogeny of the extinct marsupial wolf. Nature 340:465–467.

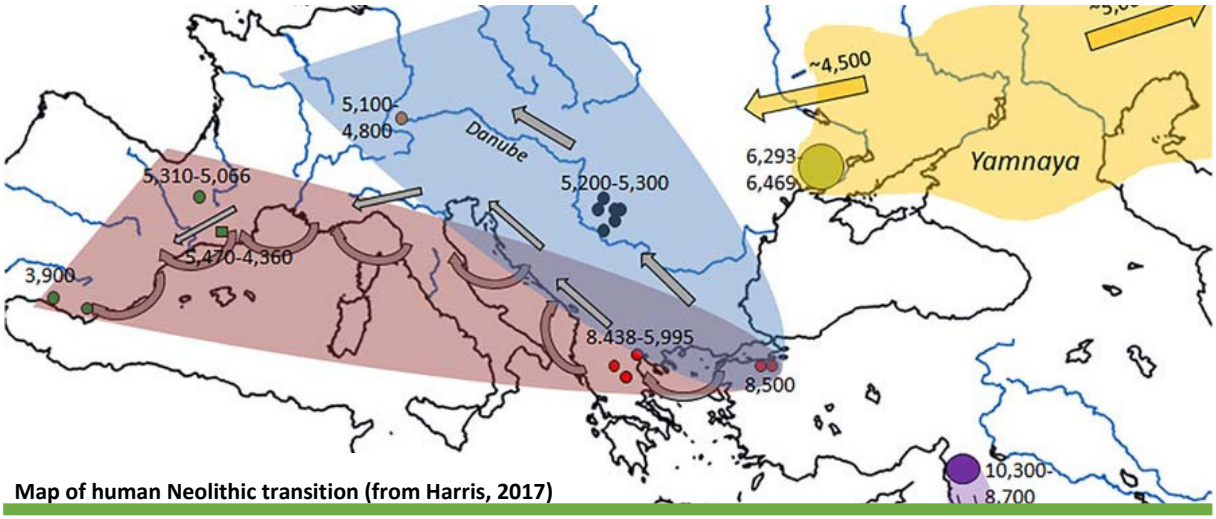
Vilstrup JT, Seguin-Orlando A, Stiller M, Ginolhac A, Raghavan M, Nielsen SCA, Weinstock J, Froese D, Vasiliev SK, Ovodov ND, et al. 2013. Mitochondrial Phylogenomics of Modern and Ancient Equids. Lalueza-Fox C, editor. PLoS One 8:e55950.

Vohr SH, Buen Abad Najar CF, Shapiro B, Green RE. 2015. A method for positive forensic identification of samples from extremely low-coverage sequence data. BMC Genomics 16:1034.

Willerslev E, Cooper A. 2005. Ancient DNA. Proc. Biol. Sci. 272:3–16.

Woodward SR, Weyand NJ, Bunnell M. 1994. DNA sequence from Cretaceous period bone fragments. Science 266:1229–32.

Yang DY, Watt K. 2005. Contamination controls when preparing archaeological remains for ancient DNA analysis. J. Archaeol. Sci. 32:331–336.



Map of human Neolithic transition (from Harris, 2017)

CHAPTER 2

The peopling of Europe: from Neolithic to the Iron Age

As mentioned above, the advances in the aDNA field, such as the improvement of sequencing methodologies and the analysis of the genomes of ancient peoples, have facilitated the determination of the genealogical relationships between humans as well as the elucidation of migration routes, diversification events and genetic admixture among various groups (Nielsen et al., 2017). Nowadays, from the analysis of the genomes of the past peoples of the European area, it was emerged that the modern-day patterns of genomic variation were in fact shaped by several important demographic events in the past. As described in Günther and Jakobsson (2016), these events include the first peopling of Europe, the Neolithic transition, and later migrations during the Bronze Age.

2.2 The Neolithic transition

The transition from a hunter-gatherer lifestyle to a sedentary farming lifestyle has been an exceptionally major change in human history, constituting, in fact, the basis for the rise of human civilizations. Thanks to the many studies carried out so far, we know that the so-called 'Neolithic transition', occurred indipendently in different areas of the world (Diamond and Bellwood, 2003; Günther and Jakobsson, 2016).

For Western Eurasia, evidences of change to a life based on agriculture and animal domestication have been found in The Fertile Crescent region of the Near East approximately 11,000 - 12,000 years ago (ya). From there, farming spread into Anatolia and Europe, reaching Scandinavia and the British Isles around 6,000 ya (Günther and Jakobsson, 2016). Archaeological investigations have also suggested that farming spread through two different

routes across the European continent: *(i)* one route along the Northern Mediterranean coast into Iberia (well represented by the Impressa and Cardial cultures), and *(ii)* the second one that going on Northwest way, into Hungary and Central Europe (well represented by the Linearbandkeramic culture (LBK)) (Harris, 2017).

The researchers have formulated several hypotheses in order to explain the transition from a hunter-gathering to a farming or domestication lifestyle (Günther and Jakobsson, 2016; Harris, 2017): *(i)* a '*demic model hypothesis'*, in which the transition was mostly produced by population movements bringing the Neolithic gene pool into Europe (Ammerman et al., 1984); *(ii)* a '*cultural hypothesis'*, based on the diffusion of agricultural ideas rather than on migrations of people (Whittle, 1996; Renfew and Boyle, 2000); and the *(iii) 'newest model'* that combine both demic and cultural diffusion, emphasizing the interactions between farmers and indigenous hunter-gatherers (HGs) (Fort, 2012).

Genetic studies carried out using uniparental markers supported both theories (Wilson et al., 2001; Chikhi et al., 2002; Haak et al., 2005; Battaglia et al., 2009; Bramanti et al., 2009; Malmstrom et al., 2009; Balaresque et al., 2010; Haak et al., 2010; Lacan et al., 2011a, 2011b; Brandt et al., 2013), but ancient genomic data from early Neolithic farmers, coming from different parts of Europe, clearly showed a strong differentiation between them and Mesolithic HGs (Gamba et al., 2014; Lazaridis et al., 2014; Skoglund et al., 2014; Haak et al., 2015; Gunther et al., 2015; Olalde et al., 2015; Cassidy et al., 2016; Hofmanovà et al., 2016). The genetic composition of the Mesolithic HGs falls outside the range of present-day Eurasians as they are genetically closer to the modern-day populations from North and Northeast of Europe. The Early Neolithic farmers seem to be genetically different from modern-day groups from the Near and Middle East, and more similar to modern-day South-Western Europeans (Günther and Jakobsson, 2016). This affinity is particularly evident for the isolated populations of Sardinians and Basques (Günther et al., 2015; Günther and Jakobsson, 2016).

The source population of the European Neolithic groups was recognized in the Early farmers from Anatolia and the Levant (Mathieson et al., 2015; Omrak et al., 2016; Lazaridis et al., 2016, Kılınç et al., 2016), in substantial agreement with the archaeological data (Childe, 1925; Günther and Jakobsson, 2016). The first farmers of central Anatolia – who lived 10 000 years BP with early pre-pottery Neolithic economical practices including small-scale cultivation, but also relying on foraging – were organized in small transition groups (Kılınç et al., 2016) constituting the genetic basis of the first expansion of farmers within Anatolia and

the Near East and then into Europe (Lazaridis et al., 2016; Kılınç et al., 2016). Instead, other groups e.g. from the eastern Fertile Crescent (Lazaridis et al., 2016; Broushaki et al., 2016; Gallego-Llorente et al., 2016) provided limited genetic material to the early European farmers (Günther and Jakobsson, 2016). The discrepancy between modern-day Anatolian genetic make-up and Neolithic groups of the area can be explained by demographic changes occurred since that period in Anatolia, such as gene-flow from the east (Omrak et al., 2016).

Regarding the models of the farming spread into the European continent, aDNA analyses on samples from different parts of Europe seems consistent (Cassidy et al., 2016; Fernández et al., 2014; Hofmanová et al., 2016) with the archaeological-based hypothesis according to which this phenomenon may have followed two different routes: one along the Danube river into central Europe; the other along the Mediterranean coastline (Özdoğan, 2011).

A substantially higher genetic diversity of farming groups (Skoglund et al., 2014; Gamba et al., 2014) can be considered as a consequence of the spread of farming practices, that probably allowed to support larger groups (Günther and Jakobsson, 2016).

Beyond to the demic and the cultural diffusion models (Pinhasi et al., 2005; Fort, 2012), the aDNA analyses on middle Neolithic farmers, showing different and increased fractions of genetic ancestry from HGs groups (Haak et al., 2015; Skoglund et al., 2014; Günther et al., 2015), demonstrate an admixture between hunter-gatherers and farmers (Günther and Jakobsson, 2016).

Although it is likely that hunter-gatherer lifestyle was replaced by farming, the farming groups assimilated hunter-gatherers giving rise to differently admixed groups across Europe, according to a pattern that can still be seen in current-day Europeans (Skoglund et al., 2014; Lazaridis et al., 2014). Furthermore, in middle Neolithic and early Chalcolithic (6000–4500 ya) farmers show additional admixture with Mesolithic hunter-gatherers if compared to early Neolithic groups (Haak et al., 2015; Günther et al., 2015). This seems to indicate that admixture between the two groups occurred in different periods and in different regions of Europe (Günther and Jakobsson, 2016). The admixture process have continued for at least two millennia which raises the question where the genetically HGs populations were settled during the Neolithic. Some researchers suggest that those groups moved to the Atlantic coast during the early Neolithic from where they re-surged later on (e.g. Haak et al., 2015; Brandt et al., 2013), but it also the hypothesis that some HGs groups still lived alongside the first farmers (Bollongino et al., 2013).

2.3 The late Neolithic and the Bronze Age

As in the early Neolithic, also during the late Neolithic and Bronze Age, genomic data support large-scale migrations with a massive impact on peoples of Europe.

The sequencing of the genome of a 24,000-year-old Siberian boy, known as 'MA-1' (Raghavan et al., 2014) showed affinities to modern-day Europeans and Native Americans, but not to East Asians. To explain this result it was supposed a scenario in which the boy belonged to an ancient north Eurasian population that likely populated much of Northeastern Eurasia until some millennia ago and that contributed to both the first Americans and Europeans (Günther and Jakobsson, 2016). Although this group is likely to have contributed genetic material to Europe for a long time the impact on central and western Europe dates to the late Neolithic (Haak et al., 2015; Allentoft et al., 2015).

The discovery of the genomic component of the Yamnaya herder populations of the Russian Pontic-Caspian Steppes in Corded Ware people in Europe (Allentoft et al., 2015; Haak et al., 2015) documents that another influx of genomic material from outside Europe occurred approximately 4,500 ya (Harris, 2017). This genetic information seems to reveal extensive admixture between European and Steppe peoples and, not a full-scale population replacement (Harris, 2017). A significantly lower Yamnaya genomic ancestry on the X chromosome compared to the autosomes in Bronze Age samples has been identified by a recent study (Goldberg et al., 2017), albeit it has been pointed out a possible bias in estimating admixture proportions (Lazaridis and Reich, 2017). Nevertheless, the decline of the Yamnaya genomic component after the first part of the Bronze Age (Haak et al., 2015), seems to indicate that the gene flow was not protracted over a very long time (Harris, 2017). According to archaeological data (Anthony, 1990, 2010), a migration from the Steppe to Europe in this period had been already hypothesized. The development of horse riding and the invention of chariots could have facilitated this phenomenon, and it is supposed to have brought a Yamnaya cultural influence and the Indo-European languages to Europe. Furthermore, this migration and the consequent admixture between the Yamnaya and the European peoples are believed to have originated the Corded Ware culture from the Late Neolithic to the Early Bronze Age (ca 4,800-4,200 ya) (Harris, 2017). A recent study (Allentoft et al., 2015) has demonstrated that genome-wide data from Corded Ware samples in Germany show the largest Yamnaya component, data from Bell Beaker in Germany an intermediate degree, while Hungarian samples (e.g., Vatya and Maros cultures) the least one. Although a close correspondence between genomic ancestry and material culture is not

necessarily expected, it has been supposed that varying degrees of influence from the Yamnaya culture could be detectable in the archaeological record as a result of variable levels of admixture (Harris, 2017).

From the late Neolithic and through the Bronze Age, both an increase in the indigenous HG component and a decline in the Yamnaya one are documented. This result seems to exclude a consistent gene flow from the Steppe, suggesting a temporally limited migration from that area (Haak, 2015). These data seem to indicate also that HGs were not replaced completely and continued to be admixed. Furthermore, it has been pointed out that the Yamnaya also spread eastward from the Steppe to the Altai region in southern Siberia. Here, genome-wide evidence from Afanasevo Culture samples shows an almost unique Yamnaya ancestral profile, which seems to strongly indicate a colonization without admixture (Allentoft et al., 2015).

It has been highlighted that, in addition to the hunter-gatherers' recolonization of Europe after the LGM, and the events linked to the 'Neolithic transition', "*the late Neolithic/Bronze Age migrations from the east are likely the third most influential event for the composition and gradients of genomic variation among modern-day Europeans*" (Günther and Jakobsson 2016).

2.3 European population makeup after the Bronze Age

Human migrations took place also after the Bronze Age. The groups involved in these processes, however, were not as highly differentiated as during the Neolithic (Lazaridis et al., 2016), when the populations showed a degree of diversity similar to that existing between modern-day continental groups (Skoglund et al., 2014). Due to this situation, the genetic composition of Europe populations was starting to look like the modern-day pattern (Günther and Jakobsson 2016). DNA analyses, both on modern and ancient samples, contributed to the reconstruction of population events of the last three thousand years (e.g. Patterson et al., 2012; Ralph and Coop, 2013; Hellenthal et al., 2014; Leslie et al., 2015; Busby et al., 2015; Martiniano et al., 2016; Schiffels et al., 2016). Several studies on the British Isles population history focused on the Iron Age (Martiniano et al., 2016; Schiffels et al., 2016), the Roman period (Leslie et al., 2015; Martiniano et al., 2016), the Anglo-Saxon period (Leslie et al., 2015; Schiffels et al., 2016) and the Viking migrations (Leslie et al., 2015). The populations of the European mainland were shaped by several small and large-scale migrations (Ralph and Coop, 2013; Hellenthal et al., 2014; Busby et al., 2015) – originating within as well as

outside Europe – which produced the isolation-by-distance pattern that we can see in modern Europeans (Menozzi et al., 1978; Lao et al., 2008; Novembre et al., 2008). It has also been suggested that the growing population size in Europe made later migrations less influential on demography since the reltive fraction of migrants was decreasing (Günther and Jakobsson, 2016).

References:

Allentoft ME, Sikora M, Sjögren K-G, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, et al. 2015. Population genomics of Bronze Age Eurasia. Nature 522:167–172.

Anthony DW, 1993. The horse, the wheel, and language : how Bronze-Age riders from the Eurasian steppes shaped the modern world.Princeton Univerity press.

Anthony DW, 2010. Migration in Archeology: The Baby and the Bathwater. Am. Anthropol. 92:895–914.

Balaresque P, Bowden GR, Adams SM, Leung H-Y, King TE, Rosser ZH, Goodwin J, Moisan J-P, Richard C, Millward A, et al. 2010. A Predominantly Neolithic Origin for European Paternal Lineages. Penny D, editor. PLoS Biol. 8:e1000285.

Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S, Marjanovic D, Primorac D, et al. 2009. Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur. J. Hum. Genet. 17:820–30.

Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkošová Z, Powell A, Burger J. 2013. 2000 Years of Parallel Societies in Stone Age Central Europe. Science (80-.). 342.

Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind C-J, et al. 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science 326:137–40. doi:10.1126/science.1176869.

Brandt G, Haak W, Adler CJ, Roth C, Szecsenyi-Nagy A, Karimnia S, Moller-Rieker S, Meller H, Ganslmeier R, Friederich S, et al. 2013. Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. Science (80-.). 342:257–261.

Broushaki F, Thomas MG, Link V, López S, van Dorp L, Kirsanow K, Hofmanová Z, Diekmann Y, Cassidy LM, Díez-del-Molino D, et al. 2016. Early Neolithic genomes from the eastern Fertile Crescent. Science (80-.). 353.

Busby GBJ, Hellenthal G, Montinaro F, Tofanelli S, Bulayeva K, Rudan I, Zemunik T, Hayward C, Toncheva D, Karachanak-Yankova S, et al. 2015. The Role of Recent Admixture in Forming the Contemporary West Eurasian Genomic Landscape.

Chikhi L, Nichols RA, Barbujani G, Beaumont MA. 2002. Y genetic data support the Neolithic demic diffusion model. Proc. Natl. Acad. Sci. U. S. A. 99:11008–13. doi:10.1073/pnas.162158799.

Diamond J, Bellwood P. 2003. Farmers and Their Languages: The First Expansions. Science (80-.). 300:597–603.

Fernández E, Pérez-Pérez A, Gamba C, Prats E, Cuesta P, Anfruns J, Molist M, Arroyo-Pardo E, Turbón D. 2014. Ancient DNA analysis of 8000 B.C. near eastern farmers supports an early neolithic pioneer maritime colonization of Mainland Europe through Cyprus and the Aegean Islands. PLoS Genet. 10:e1004401.

Fort J. 2012. Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. Proc. Natl. Acad. Sci. U. S. A. 109:18669–73.

Gallego Llorente M, Connell S, Jones ER, Merrett D, Jeon J, Eriksson A, Siska V, Gamba C, Meiklejohn C, Beyer R, et al. 2016. The genetics of an early Neolithic pastoralist from the Zagros, Iran. bioRxiv.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. 5:5257.

Goldberg A, Günther T, Rosenberg NA, Jakobsson M. 2017. Ancient X chromosomes reveal contrasting sex bias in Neolithic and Bronze Age Eurasian migrations. Proc. Natl. Acad. Sci. U. S. A. 114:2657–2662.

Günther T, Jakobsson M. 2016. Genes mirror migrations and cultures in prehistoric Europe — a population genomic perspective. Curr. Opin. Genet. Dev. 41:115–123.

Günther T, Valdiosera C, Malmström H, Ureña I, Rodriguez-Varela R, Sverrisdóttir ÓO, Daskalaki EA, Skoglund P, Naidoo T, Svensson EM, et al. 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proc. Natl. Acad. Sci. U. S. A. 112:11917–22.

Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CSI, Brandt G, Schwarz C, Nicklisch N, et al. 2010. Ancient DNA from European early neolithic farmers reveals their near eastern affinities. PLoS Biol. 8:e1000536. doi:10.1371/journal.pbio.1000536.

Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R, Heyd V, Hawkesworth C, Pike AWG, Meller H, Alt KW. 2008. Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. Proc. Natl. Acad. Sci. U. S. A. 105:18226–31.

Harris EE. 2017. Demic and cultural diffusion in prehistoric Europe in the age of ancient genomes. Evol. Anthropol. Issues, News, Rev. 26:228–241.

Hellenthal G, Busby GBJ, Band G, Wilson JF, Capelli C, Falush D, Myers S. 2014. A genetic atlas of human admixture history. Science 343:747–751.

Hofmanová Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Díez-Del-Molino D, van Dorp L, López S, Kousathanas A, Link V, et al. 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. Proc. Natl. Acad. Sci. U. S. A. 113:6886–91.

Kılınç GM, Omrak A, Özer F, Günther T, Büyükkarakaya AM, Bıçakçı E, Baird D, Dönertaş HM, Ghalichi A, Yaka R, et al. 2016. The Demographic Development of the First Farmers in Anatolia. Curr. Biol. 26:2659–2666.

Lacan M, Keyser C, Ricaut F-X, Brucato N, Duranthon F, Guilaine J, Crubézy E, Ludes B. 2011. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proc. Natl. Acad. Sci. U. S. A. 108:9788–91.

Lacan M, Keyser C, Ricaut F-X, Brucato N, Tarrús J, Bosch A, Guilaine J, Crubézy E, Ludes B. 2011. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. Proc. Natl. Acad. Sci. U. S. A. 108:18255–9.

Lao O, Lu TT, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, Balascakova M, Bertranpetit J, Bindoff LA, Comas D, et al. 2008. Correlation between Genetic and Geographic Structure in Europe. Curr. Biol. 18:1241–1248.

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. Nature 536:419–424.

Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513:409–413.

Lazaridis I, Reich D. 2017. Failure to replicate a genetic signal for sex bias in the steppe migration into central Europe. Proc. Natl. Acad. Sci. U. S. A. 114:E3873–E3874.

Leslie S, Winney B, Hellenthal G, Davison D, Boumertit A, Day T, Hutnik K, Royrvik EC, Cunliffe B, Lawson DJ, et al. 2015. The fine-scale genetic structure of the British population. Nature 519:309–314.

Malmström H, Gilbert MTP, Thomas MG, Brandström M, Storå J, Molnar P, Andersen PK, Bendixen C, Holmlund G, Götherström A, et al. 2009. Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. Curr. Biol. 19:1758–62.

Martiniano R, Caffell A, Holst M, Hunter-Mann K, Montgomery J, Müldner G, McLaughlin RL, Teasdale MD, van Rheenen W, Veldink JH, et al. 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. Nat. Commun. 7:10326.

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528:499–503.

Menozzi P, Piazza A, Cavalli-Sforza L. 1978. Synthetic maps of human gene frequencies in Europeans. Science 201:786–92.

Nielsen R, Akey JM, Jakobsson M, Pritchard JK, Tishkoff S, Willerslev E. 2017. Tracing the peopling of the world through genomics. Nature 541:302–310.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, et al. 2008. Genes mirror geography within Europe. Nature 456:98–101.

Olalde I, Schroeder H, Sandoval-Velasco M, Vinner L, Lobón I, Ramirez O, Civit S, García Borja P, Salazar-García DC, Talamo S, et al. 2015. A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. Mol. Biol. Evol. 32:3132–42.

Omrak A, Günther T, Valdiosera C, Svensson EM, Malmström H, Kiesewetter H, Aylward W, Storå J, Jakobsson M, Götherström A. 2016. Genomic Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool. Curr. Biol. 26:270–275.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics 192:1065–93.

Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford Jr TW, Orlando L, Metspalu E, et al. 2013. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature 505:87–91.

Ralph P, Coop G. 2013. The Geography of Recent Genetic Ancestry across Europe. Tyler-Smith C, editor. PLoS Biol. 11:e1001555.

Schiffels S, Haak W, Paajanen P, Llamas B, Popescu E, Loe L, Clarke R, Lyons A, Mortimer R, Sayer D, et al. 2016. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. Nat. Commun. 7:10408.

Skoglund P, Malmström H, Omrak A, Raghavan M, Valdiosera C, Günther T, Hall P, Tambets K, Parik J, Sjögren K-G, et al. 2014. Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. Science (80-.). 344.

Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. Proc. Natl. Acad. Sci. U. S. A. 98:5078–83.

Skeletal remains coming from the necropolis of Novilara (grave 171) (photo by Chiara Delpino)

CHAPTER 3

Aims of the thesis

According to recent studies, Italian population is characterized by a higher degree of internal genomic variability than other European populations (Di Gaetano et al., 2012; Brisighelli et al., 2012; Boattini et al., 2013; Sarno et al., 2014, Fiorito et al., 2016). This scenario is the result of complex demographic dynamics, dating back mainly from Late Palaeolithic and Neolithic, but also dating to Metal Ages (Piazza et al., 1988; Pesando, 2005; Boattini et al., 2013; Capocasa et al., 2016; Sazzini et al., 2016), the Middle Ages and the Early Modern Period, which have influenced in a more-or-less marked way the present-day Italian genetic pool.

In particular, an appreciable population structure for Y-chromosome lineages and a more homogeneous mitochondrial DNA background were highlighted in Italy (Capelli et al., 2007; Brisighelli et al., 2012; Boattini et al., 2013). The population dynamics that shaped the Italian gene pool are not completely clear and, among the several studies carried out on the current Italian genetic variability (Piazza et al., 1988; Barbujani et al., 1995; Boattini et al., 2013; Sarno et al., 2014; Capocasa et al., 2014; Sazzini et al., 2016; Sarno et al., 2017), only a few of them are based on ancient populations. These studies mainly focused on specimens recovered from the mainland Italy, in particular Etruscans and Lombards (Vernesi et al., 2004; Guimaraes et al., 2009; Ghirotto et al., 2013; Vai et al., 2015), or from Sardinia that, however, discloses a particular genetic history and is a well-known as an outlier in the general European genetic landscape (Caramelli et al., 2007; Modi et al., 2017; Olivieri et al., 2017).

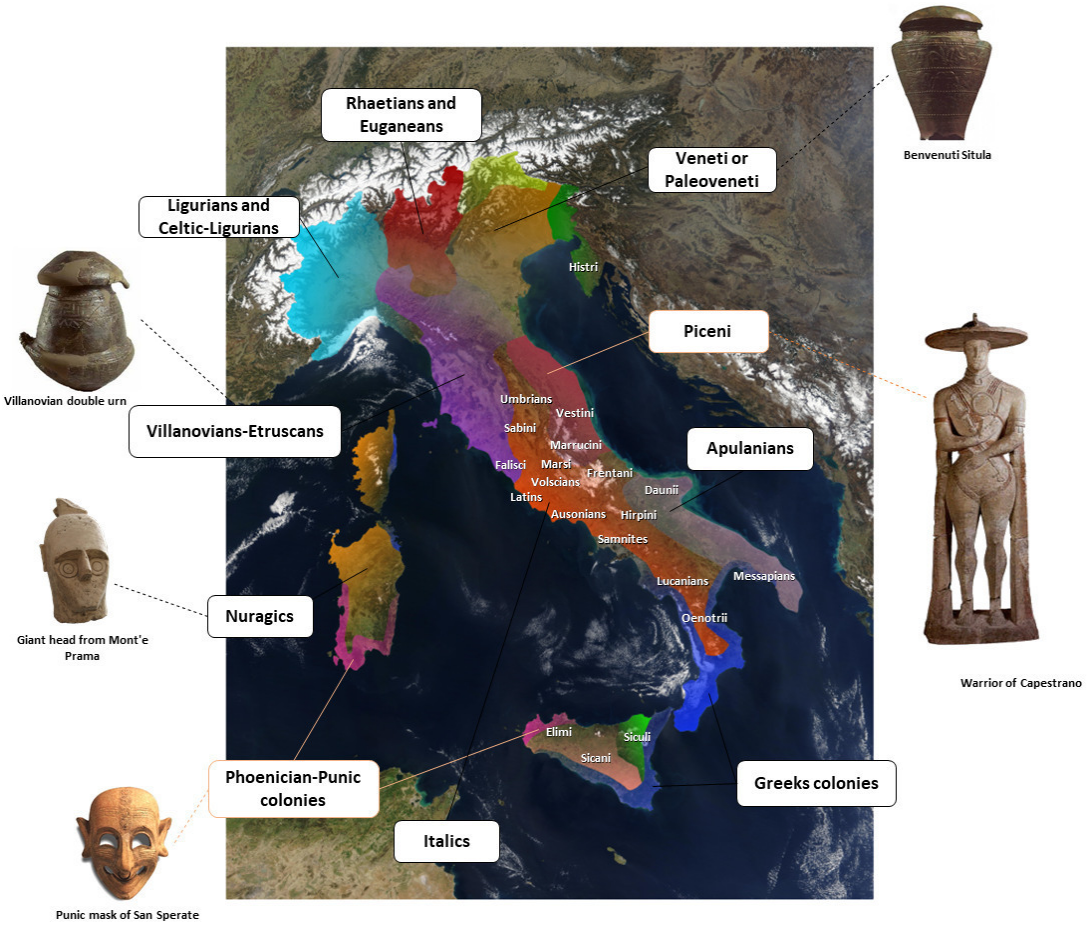


Figure 3.1 | Map of the pre-Roman peoples of Early Italy.

Before its political, juridical, linguistic, and cultural unification under the Roman Empire (Pallottino, 1991), the Italian Peninsula was a mixture of regional groups characterized by different cultural identities, languages, and dialects (Bietti Sestrieri, 2010). At the present state of research, no sufficient information is available about the origin and possible events of admixture of these populations, and our knowledge is still almost incomplete from a genetic point of view. It is an extremely complex archaeological picture characterized by different facies that go from the homogeneity that characterizes them in the Bronze Age, to a regional diversification in the Early Iron Age (Bietti Sestrieri, 2010). In this respect, a research based on the study of ancient DNA would allow a better understanding of the past events that have led to the current genetic pattern.

In traditional archaeological terms, the Italian Early Iron Age covers the 9th and 8th centuries BC and it is during this period that the first communities with strong and well defined cultural identities appeared (e.g. Etruscans, Piceni, Umbrians, Samnites) (Figure 3.1).

It was a period of great changes characterized by the adoption of new productive technologies, important social transformations, and more extended trade-routes with a full interaction both with continental Europe and Easter-Mediterranean regions (Buti and Devoto, 1974; Mallegni and Marongiu, 2010). The economy, previously based on a small-scale manufacturing, began to accommodate new classes of specialized craftsmen, thereby encouraging the growth of trades and a production on a larger scale (Bietti Sestrieri, 2010). Moreover, during all this period there were increasing contacts with the Celts, Phoenician, and Greek colonists (Devoto, 1977; Pesando, 2005).

In the wide panorama of Iron Age populations, in my PhD research, I focused the attention on the population of the Piceni from Novilara necropolis (8th-7th century BC) localized in the Marche region (PU) and on the Punic population from the 'southern necropolis' of Tharros (OR) (7th-2nd century BC) located in Sardinia and the necropolis of Lilybaeum (TP) (6th-3th century BC), both recently excavated with a rigorous stratigraphic methodology.

3.1 CASE STUDY I

Identifying the genetic legacy of the Piceni: a preliminary survey from Novilara necropolis (PU), 8th-7th century BC

Abstract: during the Iron Age, in Italy, the archaeological data provide documentary evidence of the appearance of the first communities with strong and well defined cultural identities. At the moment, only a few studies report genetic data about the Italian populations of this period and, in particular, the Piceni have never been studied from this point of view. A pilot research project, based on both genetic and archaeological approaches, has been started in the Novilara necropolis (dated at 8th-7th century BC) localized in the Marche region (central Italy). This archaeological site represents an exceptional evidence due to the presence of more than 300 graves excavated so far, characterized by an abundance of grave goods and a good conservation status of the skeletal remains. To shed light on the ancient genetic diversity of Italy of this period and to understand the contribution of the Piceni population in shaping the present-day Italian gene pool, the HVS-I region of mitochondrial DNA (mtDNA) in a first set of 27 individuals from Novilara necropolis was examined by high-coverage next-generation sequencing (NGS). Typical ancient DNA damage pattern in the analysed sequences and their comparison with those of the researchers involved in this project confirm the authenticity of the obtained data. Moreover, a forensic analysis was performed for both autosomal STRs and Indels among some individuals buried in the same pit grave to establish the possible kinship relationships.

Keywords: Italic populations, Iron Age, Piceni, ancient DNA, mitochondrial DNA, HVS-I region, Next-Generation Sequencing, kinship relationship, population genetic.

3.2. CASE STUDY II

Deciphering the identity and settlement of the "Phoenician-Punic" civilization: a first genetic study on Tharros (OR) and Lilybaeum (TP) sites

Abstract: The timing and modalities concerning the identity and expansion of the 'Phoenician' civilization and the formation and diffusion of the 'Punic' culture - linked to the Carthage cultural and territorial expansion - represent in the Phoenician-Punic studies a vexed question. In order to contribute to the reconstruction of the 'Phoenician-Punic' settlement in the central-western Mediterranean area, a research project has been started in the Tharros southern necropolis (OR, Sardinia) and Lilybaeum site (TP, Sicily) based on a multidisciplinary approach that combines the contributions of archaeological and genetic investigations *in primis*. In the present study, conducted on a first selection of bone samples, classical methods of mitochondrial DNA analysis (HVS-I and SNPs of the coding region) have been combined with new generation techniques (NGS) to obtain ancient whole genomes. This research, therefore, provides a pioneer survey in the Phoenician-Punic context, in order to define the target population and expand the knowledge on migration flows and the relationship between ancient and present-day populations of the Mediterranean area, to trace the ethnic origin, and to understand whether will be maintained a genetic continuity with those who nowadays still live in the same territories.

Keywords: Iron Age, Phoenician-Punics, Tharros, Lilybaeum, ancient DNA, mitochondrial DNA, HVS-I region, Next-Generation Sequencing, whole genome, kinship relationship, population genetic.

References:

Bietti Sestrieri AM. 2010. L'Italia nell'età del bronzo e del ferro. Dalle palafitte a Romolo (2200-700 a.C.). Carrocci Editore.

Boattini A, Martinez-Cruz B, Sarno S, Harmant C, Useli A, Sanz P, Yang-Yao D, Manry J, Ciani G, Luiselli D, et al. 2013. Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata. PLoS One 8:e65441.

Brisighelli F, Álvarez-Iglesias V, Fondevila M, Blanco-Verea A, Carracedo A, Pascali VL, Capelli C, Salas A. 2012. Uniparental markers of contemporary Italian populationreveals details on its pre-Roman heritage. Caramelli D, editor. PLoS One 7:e50794.

Buti GG and Devoto G 1974. Preistoria e storia delle regioni d'Italia: una introduzione. Sansoni università.

Capelli C, Brisighelli F, Scarnicci F, Arredi B, Caglia' A, Vetrugno G, Tofanelli S, Onofri V, Tagliabracci A, Paoli G, et al. 2007. Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. Mol. Phylogenet. Evol. 44:228–39.

Capocasa M, Anagnostou P, Bachis V, Battaggia C, Bertoncini S, Biondi G, Boattini A, Boschi I, Brisighelli F, Caló CM, et al. 2014. Linguistic, geographic and genetic isolation: a collaborative study of Italian populations. J. Anthropol. Sci. = Riv. di Antropol. JASS 92:201–31.

Caramelli D, Vernesi C, Sanna S, Sampietro L, Lari M, Castrì L, Vona G, Floris R, Francalacci P, Tykot R, et al. 2007. Genetic variation in prehistoric Sardinia. Hum. Genet. 122:327–36.

Devoto G, 1977. Gli antichi Italici. Vallecchi Editore.

Di Gaetano C, Voglino F, Guarrera S, Fiorito G, Rosa F, Di Blasio AM, Manzini P, Dianzani I, Betti M, Cusi D, et al. 2012. An overview of the genetic structure within the Italian population from genome-wide data. PLoS One 7:e43759.

Fiorito G, Di Gaetano C, Guarrera S, Rosa F, Feldman MW, Piazza A, Matullo G. 2016. The Italian genome reflects the history of Europe and the Mediterranean basin. Eur. J. Hum. Genet. 24:1056–1062.

Ghirotto S, Tassi F, Fumagalli E, Colonna V, Sandionigi A, Lari M, Vai S, Petiti E, Corti G, Rizzi E, et al. 2013. Origins and Evolution of the Etruscans' mtDNA. Hawks J, editor. PLoS One 8:e55519.

Guimaraes S, Ghirotto S, Benazzo A, Milani L, Lari M, Pilli E, Pecchioli E, Mallegni F, Lippi B, Bertoldi F, et al. 2009. Genealogical discontinuities among Etruscan, Medieval, and contemporary Tuscans. Mol. Biol. Evol. 26:2157–66.

Mallegni F and Marongiu S. 2010. Il popolamento dell'Italia:dalle origini all'età del ferro. In "Dal Bronzo al Ferro: sulla possibile origine anatolica degli Etruschi". Systema Naturae. Vol 10.

Modi A, Tassi F, Susca RR, Vai S, Rizzi E, Bellis G De, Lugliè C, Gonzalez Fortes G, Lari M, Barbujani G, et al. 2017. Complete mitochondrial sequences from Mesolithic Sardinia. Sci. Rep. 7:42869.

Olivieri A, Sidore C, Achilli A, Angius A, Posth C, Furtwängler A, Brandini S, Capodi-ferro MR, Gandini F, Zoledziewska M, et al. 2017. Mitogenome Diversity in Sardinians: a Genetic Window onto an Island's Past. Mol. Biol. Evol.

Pallottino M. 1991. A History of Earliest Italy. University of Michigan Press.

Pesando F. 2005. L'Italia antica. Culture e forme del popolamento nel I millennio a. C. Roma: Carocci editore. 326.

Piazza A, Cappello N, Olivetti E, Rendine S. 1988. A genetic history of Italy. Ann. Hum. Genet. 52, 203–213

Sarno S, Boattini A, Carta M, Ferri G, Alù M, Yao DY, Ciani G, Pettener D, Luiselli D. 2014. An Ancient Mediterranean Melting Pot: Investigating the Uniparental Genetic Structure and Population History of Sicily and Southern Italy. Caramelli D, editor. PLoS One 9:e96074.

Sarno S, Boattini A, Pagani L, Sazzini M, De Fanti S, Quagliariello A, Gnecchi Ruscone GA, Guichard E, Ciani G, Bortolini E, et al. 2017. Ancient and recent admixture layers in Sicily and Southern Italy trace multiple migration routes along the Mediterranean. Sci. Rep. 7:1984.

Sazzini M, Gnecchi Ruscone GA, Giuliani C, Sarno S, Quagliariello A, De Fanti S, Boattini A, Gentilini D, Fiorito G, Catanoso M, et al. 2016. Complex interplay between neutral and adaptive evolution shaped differential genomic background and disease susceptibility along the Italian peninsula. Sci. Rep. 6:32513.

Vai S, Ghirotto S, Pilli E, Tassi F, Lari M, Rizzi E, Matas-Lalueza L, Ramirez O, Lalueza-Fox C, Achilli A, et al. 2015. Genealogical relationships between early medieval and modern inhabitants of Piedmont. PLoS One 10:e0116801.

Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castrì L, Casoli A, et al. 2004. The Etruscans: a population-genetic study. Am. J. Hum. Genet. 74:694–704.

Warrior of Capestrano (6th century BC) (Museo Archeologico Nazionale d'Abruzzo)

CASE STUDY I

Identifying the genetic legacy of the Piceni: a preliminary survey from Novilara necropolis (PU), 8th-9th century BC

Abstract: during the Iron Age, in Italy, the archaeological data provide documentary evidence of the appearance of the first communities with strong and well defined cultural identities. At the moment, only a few studies report genetic data about the Italian populations of this period and, in particular, the Piceni have never been studied from this point of view. A pilot research project, based on both genetic and archaeological approaches, has been started in the Novilara necropolis (dated at 8th-7th century BC) localized in the Marche region (central Italy). This archaeological site represents an exceptional evidence due to the presence of more than 300 graves excavated so far, characterized by an abundance of grave goods and a good conservation status of the skeletal remains. To shed light on the ancient genetic diversity of Italy of this period and to understand the contribution of the Piceni population in shaping the present-day Italian gene pool, the HVS-I region of mitochondrial DNA (mtDNA) in a first set of 27 individuals from Novilara necropolis was examined by high-coverage next-generation sequencing (NGS). Typical ancient DNA damage pattern in the analysed sequences and their comparison with those of the researchers involved in this project confirm the authenticity of the obtained data. Moreover, a forensic analysis was performed for both autosomal STRs and Indels among some individuals buried in the same pit grave to establish the possible kinship relationships.

Keywords: Italic populations, Iron Age, Piceni, ancient DNA, mitochondrial DNA, HVS-I region, Next-Generation Sequencing, kinship relationship, population genetic.

4.1 Introduction

4.1.1 The origin of the Piceni: historical and cultural background

The traditional ethnogenesis (Strabo, Geographica 5.4.2, Pliny, Naturalis Historia 3.18.110) describes the Piceni as an ethnic group who, during the Iron Age was located along the Northern Adriatic coastline in Italy. According to Strabo and Pliny, the Piceni were

considered to have originated from a region situated North-East of Rome (Naso, 2000), the Northern area of Sabines, a region situated north-east of Rome (Naso, 2000), by means of a peculiar way of ritual migration known as the *ver sacrum* (Tagliamonte, 1994, 1996; Colonna and Tagliamonte, 1999; Naso, 2000; Antonelli, 2003; Chirassi Colombo, 2008). The hypothesis of the Sabine origin of the Piceni was also confirmed by Verrius Flaccus (*De verborum significatu*, 1st century AD), whose epitome, edited by Sextus Pompeius Festus in the 2nd or 3rd century AD (*De verborum significatu*) (Tagliamonte, 1999), was summarised in the 8th century AD by Paul the Deacon (*Excerpta ex libris Pompeii Festi de significatione verborum*). Moreover, an explicit reference to the *ver sacrum* is also found in the *Etymologiae* scholion of Isidore of Seville (Naso, 2000; Antonelli, 2003).

The "holy spring" (*ver sacrum*) is a well-known ritual in the history of pre-Roman populations that, during difficult years, vowed to sacrifice to the gods, especially Mars/Ares, everything, like animals and agricultural products, that would be born in the following spring. The children born in that period, once they reached the 21st year of age, had to leave the tribe and move to new territories. The legend tells that an animal sacred to the tribe, a woodpecker for the Piceni, *picus* in Latin, would lead these exiles to a new homeland (Antonelli, 2003). The choice of the woodpecker is not without significance in the religious landscape of pre-Roman Italy: the picus Martius, sacred to Mars in the Latin tradition, was used in the practice of augury, studying the flight of birds to interpret the will of the gods (Naso, 2000). Regarding the Piceni's augury, of great interest was also the information from Marcus Terentius Varro, the great erudite of Sabina origin, reported by Dionysius of Halicarnassus, who claimed that in the sanctuary of Mars at Tiora Matiena, an ancient Sabina religious centre, a woodpecker perched on a pole provided oracular responses. Consequently, G. Colonna has suggested, despite the silence of the sources in this regard, that the *ver sacrum* of the Piceni may have started from Tiora Matiena. This sanctuary, archaeologically unknown, should be sought in the heart of the inner Sabina, on the Apennine Mountains, as suggested by F. Ribezzo who associated the name with the Teora locality, near to Amiterno, in the present Aquila province. From this point of view, the *ver sacrum* of the Piceni could have followed the natural itinerary that from the Aquilana Mountain, through Montereale and Amatrice, leads to Ascoli, which was therefore considered to be the main centre of the new population (*caput gentis*), destined to preserve this feature for a long time (Naso, 2000).

Both the chronological and the geographical boundaries of the Piceni have been discussed for a long time. Based on archaeological data, this civilization developed from the 9th century

BC to the beginning of the 3rd century BC, when Romans conquered this territory. The Superintendent D. Lollini of the Marche region, established a cultural sequence of Piceni civilization into seven main phases: Piceno I (900-800 BC); Piceno II (800-700 BC); Piceno III (700-580 BC); Piceno IVa (580-520 BC); Piceno IVb (520-470 BC); Piceno V (470-385 BC); Piceno VI (385-268 BC) (Lollini, 1976, Naso, 2000).

According to the scheme proposed by D. Lollini, the 9th and 8th century BC, corresponding to "Piceno I" and "Piceno II" respectively, were considered as the periods of formation and consolidation of the Piceni civilization before the cultural transformation that took place with the orientalising period in the 7th century BC (Piceno III). The archaeological evidence shows that the Piceni were present in most of the whole present-day Marche region. In their whole distribution area, the Piceni shared a common culture with some distinctive local features, favoured by the peculiar landscape of Marche, mostly consisting of hills with a linear pattern and narrow river valleys extending from the Apennines to the Adriatic coast. The lack of well-defined single urban centres in this territory maintained this Adriatic culture at a proto-urban stage and, more importantly, facilitated the permanence of local cultural diversities until the dominance of Rome. Knowledge of this civilization is mostly based on archaeological findings, largely coming from necropolises (Naso, 2000).

The Piceni's burials are usually found near the settlements, and in some cases (Novilara, Montedoro-Scapezzano of Senigallia, Matelica, Ancona and Camerano) above necropolis and/or settlements of the previous periods. About the burial practices, they are characterized by use of inhumation, with the deposition of the crouched body on the right side in a simple pit grave, even if in the oldest burials some cremation depositions persisted, in continuity with the Bronze Age habits (i.e. some burials of the 9th century BC in Ancona, Numana and Matelica). In the beginning, the grave goods were absent or very scarce, while during the "Piceno II" they became richer. In addition, in this phase, the Adriatic coast became an ideal place for commercial exchanges of iron and amber, and thus several proto-urban centres were established in the area: the necropolis of Novilara fits in this context.

4.2. Materials and Methods

4.2.1. Archaeological Context

One of the most important centres of the Iron Age of the Northern Picenum area is the hill of Novilara (43° 51′ N - 12° 55′ E) (Figure 4.2.1.1) in the Province of Pesaro and Urbino. In

fact, the Novilara site is characterized by an abundance of grave goods, systematically excavated starting from the end of the 19th century, good conservation status of the skeletal remains and, lastly, by a considerable amount of scholarly literature (Brizio, 1895; Beinhauer, 1985; Delpino et al., 2016, in press).

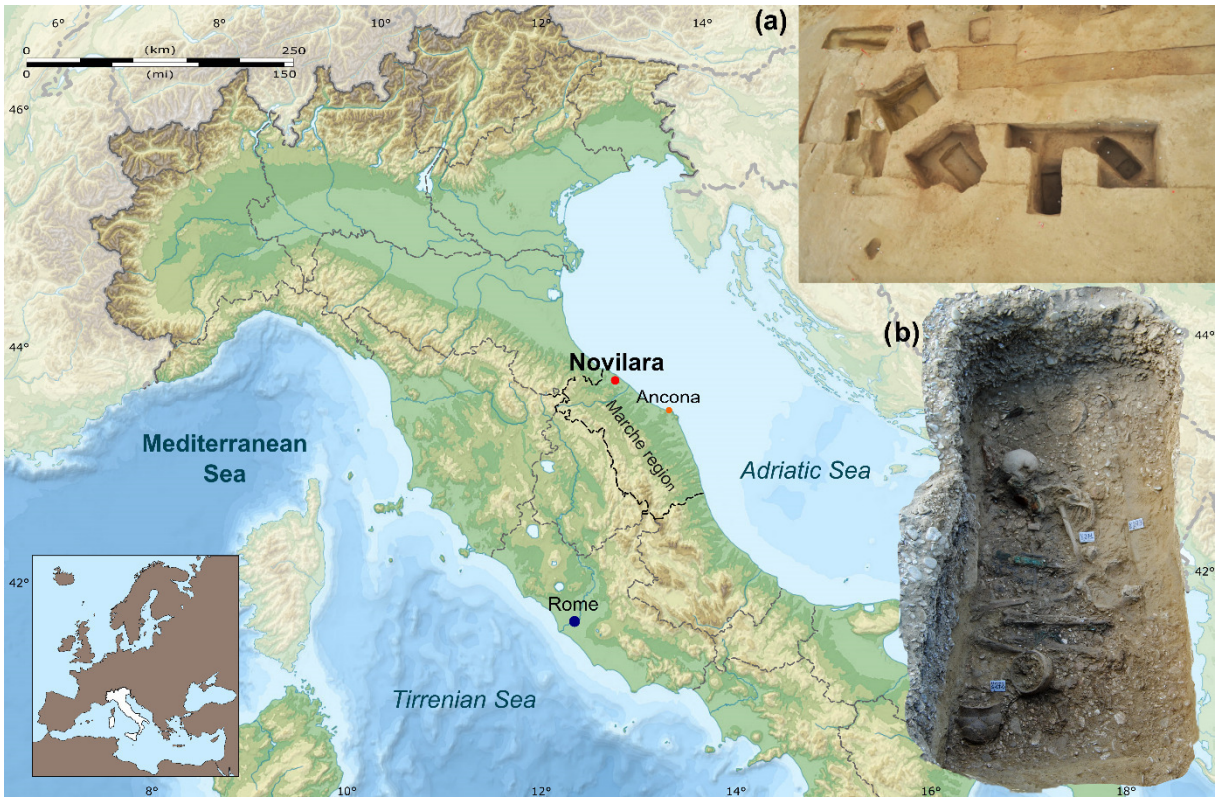


Figure 4.2.1.1 | Geographic location of the Novilara site: a) archaeological fieldworks area during the campaign of 2012, b) example of burial found during 2012-2013 (woman's grave n. 10, 7th century BC).

As mentioned above, the necropolis of Novilara has been known in literature since the late 19th century. The first excavations, in an almost clandestine way, were carried out in the1873s by the Count Bonamini in the so-called “Servici area”. Afterwards, an archaeological survey was practised in the 1891s by Ciro Antaldi (curator at the Oliveriano Archaeological Museum of Pesaro, Italy), Eugen Ludwig Bormann (German-Austrian historian), and the archaeologist Gian Francesco Gamurrini, who published the discovery in the magazine "Notizie degli Scavi". Later, between the 1892s and the 1893s, Edoardo Brizio conducted a systematic archaeological investigation in two different areas: the so-called “Molaroni area” and the “Servici area”. He discovered 263 burials with related grave goods (dating back to the 8th and 7th centuries BC), then exposed at the Oliveriano Archaeological Museum (Brizio, 1895; Beinhauer, 1985). The archaeological investigations were resumed in

the 1912s by Innocenzo Dall'Osso (Delpino et al., 2016), but all the grave goods of the 30 burials found, transported to the National Archaeological Museum of the Marches (Ancona, Italy), were forever lost in the 1944s, during the aerial bombardment that struck the city of Ancona.

In the 2006s a new research began, because of some major works related to the enlargement of a motorway. The Archaeological Superintendence of the Marche region highlighted how it was likely to recover interesting stratigraphic units close to the tunnel entrance of the new lane of the A14 highway. Thus, the Superintendence required that all the excavation procedures were followed by a group of archaeologists, and that, in case of interesting archaeological findings, the regulation of the"article 90 of the d.lsg. 42/2004" would be applied. In October 2011, the presence of some filled pits - possible graves - was noted. The archaeological investigations began in March 2012 with a small group of archaeologists and anthropologist under the supervision of the archaeologist Chiara Delpino. The excavation focused both on the northern and southern slope of the Novilara hill, for a total area of 12,550 m^2. Since the works for the enlargement to three lanes of the highway have never been stopped, the excavation strategies were determined with the "Autostrade S.p.a." company, according to the construction necessities. Thus, the entire area was divided into excavation districts, alternatively parallel or perpendicular to the entrance of the tunnel, and this approach permitted both to never stop the enlargement works and to allow the application of a rigorous excavation methodology. Later, the area was divided into 5x5 meters excavation squares, named with letters and numbers, in order to easily determine the spatial collocation of the archaeological assemblages recovered.

During the archaeological excavations of the 2012s and the 2013s, 150 new burials have been discovered, dating back to the 8th and 7th centuries BC (Figure 4.2.1.2). The burials were generally in a good conservation status. Just in restricted areas, where a more important natural erosion phenomenon took place, profound ploughings had already brought to the surface some skeletal remains and grave goods. The burials showed unitary burial practices: all the skeletons recovered were crouched on the right side with flexed legs (Delpino et al., 2016). Only the arms and the rotation of the chest showed some degree of flexibility, but it is hard to determine if this was the consequence of a deliberate behaviour or solely a random effect. The mandible was almost always connected to the cranium and in several situations, the mouth was tightened. This fact could be due to the presence of bandages, as the orderly aspect of feet seems to indicate as well, that the feet, in fact, were often placed side by side or

one over the other (Delpino et al., 2016). The burial site consisted of pits almost always occupied by a wooden structure with a lid that, decomposing, left in the ground some blackish charcoal lenses (meaning some minor thin lines of deposit), reddish secondary clays or whitish lenses. Some pits were even filled, on the top of the lid, with some marine gravel (Delpino et al., 2016).

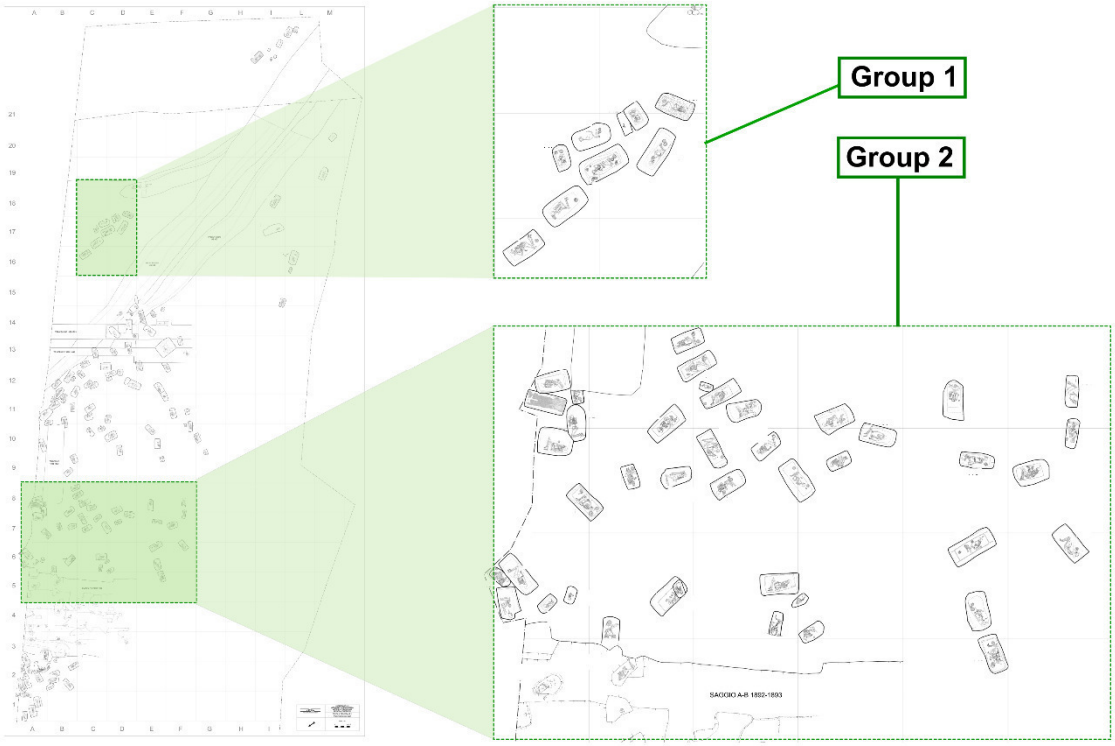


Figure 4.2.1.2 | Map of the Novilara necropolis. Within the necropolis, some areas clearly appear to have been more densely used and others show small, scattered groups of graves (for example the "Group 1" and the "Group 2").

In all the burials, with a slight preference for the ones reserved to female individuals, some grave goods were found next to or above the body, witnessing a well-defined burial practice. The simpler and more ancient goods were composed by few personal bronze ornaments, such as fibulae of different measures and types, some fictile (clay) objects, like small two-handled cups, *kothons* and *ollettes*. The more complex grave goods consisted of several personal objects, as utensils, pottery, ornaments, and clothes. Making a distinction between male and female burials, these latter ones were better structured and richer. The male burials mainly contained bronze and iron weapons and razors, while the female one's contained pottery, spools and sewing weights, personal ornaments such necklaces, fibulae, brooches and rings

often decorated with amber stones and clothes. Obviously, not all the grave goods were so rich, defining a certain degree of social structure within the community.

Among all the skeletal remains of which we determined the age, 80% belonged to adults, while 20% belonged to infants, children, and adolescences (Delpino, in press). Within the necropolis, some areas clearly appear to have been more densely used and others show small, scattered groups of graves (Figure 4.2.1.2). From an archaeological perspective, the burials repartition potentially shows that the necropolis' funerary space was structured following family or clan-related groups. For example, the "Group 1", one of the two burial clusters here presented (Figure 4.2.1.2), is clearly segregated from the other. The presence of a homogeneous number of graves of men, women, and children of all ages (newborns, infants and children), buried with wealthy grave goods dating to the middle of the 7th century BC, seem to indicate the existence in this period of familiar units of an elevated social status (Delpino, in press). In fact, one of the wealthiest infant burials found in the necropolis (burial 58) is part of this group. Moreover, the discovery of the grave (burial 57) of a young woman buried with a peculiar funerary practice (her grave pit was lining with marine pebbles) may suggest a female exogamy, yet to be tested by genetic data. In the "Group 2" here presented (Figure 4.2.1.2), the burials are distributed following a semi-circular pattern that resembles the grave mounds of other Italian Iron Age cultures, settled in central Italy from Tuscany to Abruzzo regions (Dalla Fina, 2015). In addition, in this graves group, both adults and infant were buried in equal extent.

4.2.2. Samples for genetic analysis

The human skeletal remains of 27 individuals from the Iron Age necropolis of Novilara (Table 4.2.2.1) were selected on the basis of the archaeological information and on the state of preservation of the samples. All the graves have been dated to the 8th–7th century BC by means of the stratigraphic sequence coupled with the recovered grave goods and the general historical reconstruction of the site.

The sampling of teeth and petrous bones was carried out throughout the archaeological campaign in 2011-2012 field season with all the necessary precautions described in Fortea et al. (2008) to minimize the occurrence of contamination from modern human DNA: *(i)* the samples were collected using disposable lab coat, sterile gloves, face mask, over-shoes and selected based on their status of preservation (without visible signs of damage or mineral diagenetic alteration). *(ii)* All the instruments used were decontaminated with 5% NaClO,

Ethanol and DNA-ExitusPlus™ solution (AppliChem GmbH, Darmstadt, Germany) before and after each sampling. *(iii)* Whenever possible, two samples were collected per individual and then all the specimens were delivered to the laboratory, stored in specific plastic bags, annotated with sample description, location, number of burial and stratigraphic unit. *(iv)* Buccal swab samples from all personnel involved in this study (archaeologists, anthropologists, and laboratory researchers) were taken to monitor potential sources of contamination.

Sample ID	*Material*	*Sample info*	*Century*	*Years of excavation*	*Researchers[1]*
NOR3a	Petrous bone, tooth	Grave 171	8th-7th c. BC	2012-2013	M1, M2, M3
NOR3b	Petrous bone, tooth	Grave 171	8th-7th c. BC	2012-2013	M1, M2, M3
NOR8	Petrous bones	Grave 67	8th-7th c. BC	2012-2013	M2, M4
NOR9	Petrous bone	Grave 58	8th-7th c. BC	2012-2013	M1, M2, M3
NOR10a	Petrous bone	Grave 155	8th-7th c. BC	2012-2013	M2, M4
NOR10b	Petrous bones	Grave 155	8th-7th c. BC	2012-2013	M1, M2, M4
NOR11	Petrous bone	Grave 25	8th-7th c. BC	2012-2013	M1, M3
NO12	Tooth	Grave 85	8th-7th c. BC	2012-2013	M2, M4
NO13	Tooth	Grave 89	8th-7th c. BC	2012-2013	M2, M3
NO14	Tooth	Grave 41	8th-7th c. BC	2012-2013	M1, M2, M3
NO15	Tooth	Grave 45	8th-7th c. BC	2012-2013	M1, M2, M4
NO16	Tooth	Grave 57	8th-7th c. BC	2012-2013	M1, M2, M3
NO17	Tooth	Grave 39	8th-7th c. BC	2012-2013	M2, M4
NO18	Tooth	Grave 27	8th-7th c. BC	2012-2013	M1, M2, M4
NO19	Tooth	Grave 125	8th-7th c. BC	2012-2013	M2, M4
NO20	Tooth	Grave 83	8th-7th c. BC	2012-2013	M2, M3
NO21	Tooth	Grave 128	8th-7th c. BC	2012-2013	M2, M4
NO22	Tooth	Grave 160	8th-7th c. BC	2012-2013	M2, M3
NO23	Tooth	Grave 161	8th-7th c. BC	2012-2013	M2, M3
NO24	Tooth	Grave 137	8th-7th c. BC	2012-2013	M2, M3, M4
NO25	Tooth	Grave 146	8th-7th c. BC	2012-2013	M2, M3
NO26	Tooth	Grave 87	8th-7th c. BC	2012-2013	M2, M3
NO27	Tooth	Grave 101	8th-7th c. BC	2012-2013	M2, M3, M4
NO28	Tooth	Grave 154	8th-7th c. BC	2012-2013	M2, M3
NO29	Tooth	Grave 175	8th-7th c. BC	2012-2013	M1, M2, M4
NO30	Tooth	Grave 90	8th-7th c. BC	2012-2013	M1, M2, M3
NO31	Tooth	Grave 152	8th-7th c. BC	2012-2013	M2, M3

Table 4.2.2.1 | Information about the Novilara samples. [1]researchers who have been in contact with the ancient samples during the archaeological excavation and the laboratory work.

4.2.3. Ancient DNA procedures

DNA extractions and PCR setup were performed in physically isolated work areas dedicated to ancient DNA analysis at the Laboratories of Physical Anthropology and Ancient DNA, Department of Cultural Heritage (DBC), University of Bologna, according to rigorous aDNA standards to avoid contaminations (Cooper and Poinar, 2000; Fulton et al., 2012; Knapp et al., 2012, 2015). Suitable disposable clothing (coverall suit, double pair of gloves,

over-shoes, face mask and plastic face shield) were worn during the handling and extraction of materials. The worktop and the instruments were regularly wiped with 5% commercial NaClO, 96% ethanol, DNA-ExitusPlus™ solution (AppliChem GmbH, Darmstadt, Germany) (Esser et al., 2006) after each experiment. In addition, the ancient DNA laboratory was exposed to ultraviolet radiation (λ = 245 nm) overnight for approximately 4 hours, in order to degrade the DNA on laboratory surface and equipment. Sterile materials and dedicated pipettes with aerosol resistant tips were used at each step of work. In addition, all the reagents were screened for modern DNA and stored in small volume aliquots before use. Multiple blank extractions were processed in parallel and negative controls were included in all reactions. The HVS-I sequences of the personnel involved in this study (archaeologists, anthropologists, and laboratory researchers) were compared with the genetic profiles obtained from the ancient specimens to make sure of the absence of modern contamination. PCR and post-PCR laboratory procedures (libraries preparation of amplicons and next-generation sequencing) were carried out in a separate building at the Laboratory of Molecular Anthropology and at the Centre for Genome Biology, Department of Biological, Geological and Environmental Sciences (BiGeA), University of Bologna.

4.2.4. Sample preparation

The samples were decontaminated by removing the surface layer with a Dremel® drill and then irradiated with ultraviolet (λ = 254 nm) light for 15-30 min on each side.

The tooth was cut transversally at the cementum-enamel junction before sampling dentine powder with a diamond drill bit set at low speed to avoid heating. While from the petrous bone the denser inner part was taken with a Dremel® drill and was ground into a fine powder with a mortar, as recent literature suggested that this part can provide higher endogenous DNA (Gamba et al., 2014; Pinhasi et al., 2015; Henrik et al., 2017).

Tooth and petrous bone powder were stored at 5°C until further use. When more samples were prepared on the same day, all the tools have been carefully cleaned between every sampling procedure to prevent cross-contamination.

4.2.5. Ancient DNA extraction

Between 150 to 300 mg of bone powder was used for each DNA isolation, performed by means of a silica-based method (Dabney et al., 2013) with a few modifications. This method of extraction exploits the presence of a siliceous resin inside the test tubes, that is able to

absorb nucleic acids on its surface. This happens by adding a chaotropic salt, that destroy hydrogen bonds and denature proteins. This method joins effectiveness of chromatography and the speed of centrifugation, that promote the passage of a liquid through the silica membrane. The silica absorbs almost the 90-95% of the DNA present in the solution, allowing to discard the contaminants by subsequent washing steps and centrifugations. The powdered samples were usually made in solution with an extraction solution and charged in the silica column, where the nucleic acids are selectively absorbed from the membrane around a pH ~7.5. Washing buffers and centrifugations were used to discard everything not absorbed by the membrane. At the end, the DNA was eluted by means of a proper elution buffer in basic conditions and low concentrations of salts (Tagliabracci, 2009; Caramelli, 2009). A negative control (blank extraction), consisting in 1 mL of extraction solution without a bone powder was included and treated as a regular sample, in order to monitor a possible contamination in this step of the process, either in the chemicals used or caused by the operators.

In detail, 150 to 300 mg of bone powder was decalcified and digested for 24 h with 1 mL of extraction buffer (Table 4.2.5.1) under constant agitation on a rotary mixer at 37°C.

Reagent	*Volume for 1 sample (µL)*
0.45M EDTA, pH 8.0	900
0.25 mg/mL Proteinase K	12,5
Nuclease-free water	87,5
Total	1000

Table 4.2.5.1 | Extraction buffer.

After that, samples were centrifuged at maximum speed and the supernatant was removed and transferred in a new 15-mL falcon tube with 10 mL of binding buffer (Table 4.2.5.2). A binding apparatus was constructed by forcefully fitting an extension reservoir removed from a Zymo-Spin™ V column (Zymo Research -Irvine, CA, USA), inserted into a MinElute silica spin column (Qiagen GmbH, Hilden, Germany) and then placed into a 50-mL falcon tube. To avoid contamination the Zymo-Spin™ V extension reservoir was previously submerged in a 5% commercial NaClO bath for 20 minutes, washed with nuclease-free water, and UV irradiated before use. Moreover, the binding apparatus was tested in the centrifuge for 4 minutes at 1,500 g before the use. The solution containing the binding buffer and the extraction supernatant was then poured into the extension reservoir apparatus, which was

centrifuged (1,500g for 4 minutes), rotated at 90°, and centrifuged again at 1,500 g for 2 minutes.

Reagent	*Volume for 1 sample*
5M Guanidine Hydrochloride	4.7765 gr
90mM Sodium Acetate	4000 µL
40% Isopropanol	300 µL
0,05% Tween-20	5 µL
Nuclease-free water	6193 µL
Total	10 mL

Table 4.2.5.2 | Binding buffer.

Afterwards, the silica membrane was centrifuged and washed twice with 750 µL of PE buffer (Qiagen GmbH, Hilden, Germany) at 6,000 rpm for 30 seconds and, after putting the column in a new 1.5 ml tube, the DNA was eluted by adding 35 µL of TET buffer (10mM Tris-HCL 10mM, 1mM EDTA, 0.05% Tween-20, and nuclease-free water) directly on the silica membrane. The column was incubated at room temperature for 5 minutes and finally centrifuged at 8,000 rpm for 1 minute and then at 10,000 rpm for 2 minutes. After that, 1 µL of eluted DNA was used for DNA quantification with Qubit® dsDNA HS Assay Kit (Invitrogen™ Life Technologies - Carlsbad, CA, USA), following the manufacturer's instructions.

4.2.6. mtDNA amplification

The mtDNA HVS-I control region was amplified as three overlapping fragments using primer pairs which yielded products of 179, 197 and 156 bp (L15995-H16132, L16107-H16261, L16247-H16402) (Caramelli et al., 2003) (Table 4.2.6.1), in order to obtain 360 bp, spanning from nucleotide position (np) 16024 to np 16383.

PCR reaction was performed in a final volume of 25 µL of reaction mix containing (Table). The temperature profile of the thermal cycler program was as follows: initial denaturation at 94°C for 10 minutes, 40 cycles of 45 seconds at 94°C, 1 minutes at 53°C and 1 minute at 72°C, followed by a final extension for 10 minutes at 72°C.

If the PCR reactions failed for some samples, the reagents volumes were modified: the quantity of the $MgCl_2$ was raised to increase polymerase activity. The amplification of each fragment was carried out in independent PCR reactions.

Reagents	*Volume for 1 sample (µl)*
1 X AmpliTaq Gold-Buffer	2.5
2.5 U AmpliTaq Gold DNA Polymerase	0.5
0.25 mM dNTP mix	0.625
2 mM MgCl2	2
0.2 mM primer Fwd	0.5
0.2 mM primer Rvs	0.5
0.8 mg/mL BSA	1
Nuclease-free water	15,375
DNA	2
Total	25

Table 4.2.6.1 | PCR reaction mix.

To verify the repeatability of the mitochondrial data and to confirm the authenticity of the results, for a set of randomly selected individuals, the whole experiment, from DNA extraction to DNA amplification, was performed twice (Hervella et al., 2015; Lorkiewicz et al., 2015), starting from different anatomical elements and by different researchers.

Furthermore, all DNA extracts were screened to test their appropriate molecular behaviour (Cooper and Poinar, 2000) with L15996-H16401 primers pairs (Vigilant et al., 1989) (Table 4.2.6.2), which amplify a larger fragment (~400bp), to detect possible contaminations, given that ancient DNA molecules are often fragmented to very short pieces encompassed between 60 and 150 bp (Prüfer et al., 2010; Sawyer et al., 2012). The temperature profile of the thermal cycler program was as follows: initial denaturation step at 95°C for 10 minutes, 35 cycles consisting of 45 seconds at 94°C, 1 minute at 60°C, 1 minute at 72°C, followed by a final extension for 10 minutes at 72°C.

HVS-I	*Region*	*Sequence*	*Reference*
I Fragment	L15995 H16132	5'-CCACCATTAGCACCCAAAG-3'	Caramelli et al., 2003
		5'-CTACAGGTGGTCAAGTATTTATGGT-3'	Caramelli et al., 2003
II Fragment	L16107 H16261	5'-CGCTATGTATTTCGTACATTACTGC-3'	Caramelli et al., 2003
		5'-TGGTATCCTAGTGGGTGAGG-3'	Caramelli et al., 2003
III Fragment	L16247 H16402	5'-CAACTATCACACATCAACTGCAA-3'	Caramelli et al., 2003
		5'-GATTTCACGGAGGATGGT-3'	Caramelli et al., 2003
Long Fragment	L15996 H16401	5'-CTCCACCATTAGCACCCAAAGC-3'	Vigilant et al., 1989
		5'-TGATTTCACGGAGGATGGTG-6'	Vigilant et al., 1989

Table 4.2.6.2 | PCR primer pairs used for amplification of HVS-I sequences.

4.2.7. Horizontal gel electrophoresis and purification of DNA amplicons

All PCR products and both extraction and amplification controls were visually examined by standing 1.5% agarose TBE gel, which had undergone electrophoresis, with GelRed™ Nucleic Acid Gel Stain (Biotium Inc., Fremont, USA). The amplicon products (three fragments for each sample), which produced a visible band on the electrophoresis gel, were purified using two different approaches. The majority of them were purified with the QIAquick® PCR Purification Kit (Qiagen GmbH, Hilden, Germany), following manufacturer's recommendation. This protocol is designed to purify double-stranded DNA fragments from PCR reactions resulting in high end-concentrations of DNA. Briefly, in the presence of 5X volume of PB buffer (Qiagen GmbH, Hilden, Germany) the DNA was adsorbed to the silica membrane of the MinElute spin column, while possible inhibitors (i.e. primers, nucleotides, polymerases, and salts) were discarded during a centrifugation step at 13,000 rpm for 1 minute. All the impurities remained into the silica column were washed away by means of 750 µL of PE buffer (Qiagen GmbH, Hilden, Germany) and several centrifugation passages at 13,000 rpm for 1 minute. The purified DNA amplicon was eluted with 10 µL of nuclease-free water added directly to the silica membrane.

A second procedure was used for those fragments that not showed a clear band in the electrophoresis gel, probably due to a phenomenon described as "jumping PCR" (Pääbo, 1989). An E-Gel® Electrophoresis System (Invitrogen™ Life Technologies - Carlsbad, CA) was used with a Size Select agarose gel on the iBase™ Power System (Invitrogen™ Life Technologies - Carlsbad, CA) in order to select the amplicons of the size of interest. In detail, the wells of the upper row were loaded with 22 µL of PCR product, 10 µL of 50 bp DNA Ladder (E-Gel 50 bp DNA Ladder, Invitrogen™ Life Technologies - Carlsbad, CA) was added into the small middle well (line M) and 25 µL of nuclease-free water into any remaining empty wells. The wells of the lower row were loaded with 25 µL of deionized water, except for the well of lane M (10 µL). The run was monitored periodically, so when the ladder band of the size of interest reached the reference line, the run was stopped, and the collection wells were refilled with 10 µL of nuclease-free water. Then, when the amplicon band of the expected size migrated into the well, it was collected using a pipette, paying attention not to pierce the bottom of the well. In those cases, when the bands seemed to be had overshot the collection well, the Reverse E-Gel program was used to run the band back into the collection well. Since 22 µL of sample were loaded and 25 µL were collected, diluted in 10 µL of water, a control step of quantification with Qubit® 2.0 was performed.

4.2.8. DNA library preparation

The purified PCR products were converted into blunt-end sequencing libraries using the Ion Plus Fragment Library Kit (Life Technologies, Carlsbad, CA, USA), according to manufacturer's instructions. All DNA fragments were blunt-ended, after a purification step, and then the P1 adapters belonging to the Ion Xpress Barcode Adapters 1-96 Kit (Life Technologies, Carlsbad, CA, USA) were attached to the fragments. After a second purification step, the adapters were filled in without the following purification. Both extraction negative control and library negative control were included and processed like normal samples, to monitor the absence of contamination during the experimental steps. At first, all the purified PCR products were quantified using the Qubit® dsDNA HS Assay Kit (Invitrogen™ Life Technologies - Carlsbad, CA, USA) and normalized to a quantity of 100 ng (33 ng for each fragment) to make sequencing coverage as uniform as possible within the sample.

4.2.8.1. Blunt-end repair

PCR products were pooled and blunt end-repaired using the Ion Plus Fragment Library Kit (Life Technologies, Carlsbad, CA, USA), to produce flat ending fragments where adapters would then be attacked: 100 µL of the blunt end-repair mix (Table 4.2.8.1.1) were incubated at room temperature for 20 minutes.

Reagent	*Volume for 1 reaction (µl)*
Pooled amplicons, 10-100 ng	79
5X End Repair Buffer	20
End Repair Enzyme	1
Total	100

Table 4.2.8.1.1 | Blunt-end repair.

In order to remove the residuals of the reagents of the previous reaction, a purification step with the Agencourt® AMPure® XP magnetic beads (Beckman Coulter, Brea, CA, USA). was performed. This procedure consists in addition of 180 µL of Agencourt® AMPure® XP reagent to the sample, mixing thoroughly the bead suspension with the DNA by pipetting up and down five times. After spinning and an incubation step at room temperature for 5 minutes, during which the DNA fragments bind to the magnetic beads, the sample was pulse-centrifuged and placed in a magnetic rack until the solution cleared (about 3 minutes). The

supernatant was removed and then, with the sample still in the magnetic rack, two purification steps were performed by two washes with 500 µL of freshly prepared 70% ethanol. During these washes the sample's tube was turned for four times by 180° rotations, to let the beads migrating from one side of the tube to the other, moving through the ethanol and thus getting cleaned. After each wash the supernatant was removed and after the latter one particular caution was used to completely removed residual ethanol that could inhibit further steps. The beads were then air-dried at room temperature and then the sample was removed from the magnetic rack and eluted in 25 µL of Low TE buffer. The solution was carefully mixed, and the sample was placed in the magnetic rack again. When the solution cleared, the supernatant with the purified DNA was transferred to a new 0.2 mL tube.

4.2.8.2. Adapter ligation and nick repair

Barcoded and P1 adapters belonging to the Ion Xpress Barcode Adapters 1-96 Kit (Life Technologies, Carlsbad, CA, USA) were ligated onto the 5' and 3' ends of the produced fragments by nick translation using Ion Plus Fragment Library Kit (Life Technologies, Carlsbad, CA, USA). The mixture (Table 4.2.8.2.1) contains both ligase buffer for adapter ligation, dNTPs and a nick repair polymerase, as the adapters need blunt-ended fragments, but they ligate just to one strand and thus the nick created on the other strand needs to be filled in.

Reagent	*Volume for 1 reaction (µL)*
Blunt-ended DNA	~25
10 X Ligase Buffer	10
Ion P1 Adapter	2
Ion Xpress™ Barcode X	2
dNTP mix	2
nuclease-free water	49
DNA Ligase	2
Nick Repair polymerase	8
Total	100

Table 4.2.8.2.1 | Adapter ligation. X = chosen barcode.

The temperature profile of the thermal cycler program was as follows: 25° C for 15 minutes to let the enzymes work, 72° C to inactive the enzymes, and 4° C up to 1 hour as a holding stage.

The obtained libraries were purified using 150 µL of Agencourt® AMPure® magnetic beads (Beckman Coulter, Brea, CA, USA), that is indicated for library size of 100-150 base read, in order to achieve the final volume of 20 µL.

4.2.9. Real-Time PCR

Afterwards, libraries concentrations were determined using quantitative real-time PCR (qPCR) on a 7500 Fast System (Applied BioSystems, Foster City, USA) with the Ion Library Quantitation Kit (Life Technologies, Carlsbad, CA, USA) following manufacturer's recommendation. For the qPCR, serial dilutions of the E. coli DH10B Control Library were prepared with a factor of 1:10 and 1:10000 (Table 4.2.9.1).

Standard	*Control Library*	*Nuclease-free water*	*Dilution factor*	*Concentration*
1	5 µL undiluted Control Library	45 µL	1:10	6.8 pM
2	5 µL Std 1	45 µL	1:100	0.68 pM
3	5 µL Std 2	45 µL	1:1000	0.068 pM
4	5 µL Std 3	45 µL	1:10000	0.0068 pM

Table 4.2.9.1 | Standard dilutions.

Then, the sample libraries were diluted with a 1:100 dilution factor, which means adding 2 µL of library input into 198 µL of nuclease-free water. In each well were pipetted 15 µL of qPCR reaction mix (Table 4.2.9.2), and then 5 µL of either Standard 1 - 4 or diluted libraries. Each diluted library was quantified twice, and so it was added in two wells. At last, the plate was sealed and briefly centrifuged and finally placed into the instruments.

Reagent	*Volume for 20 µL reaction*
TaqManR Fast Universal PCR Master Mix	10
20X Ion Library TaqManR Quantitation Assay	1
Nuclease-free water	4

Table 4.2.9.2 | qPCR reaction mix.

The cycle conditions consisted of 50°C for 2 minutes, 95°C for 20 seconds and 40 cycles of 3 seconds at 94°C and 30 seconds at 60°C. All barcoded libraries were diluted and pooled to obtain a final concentration of 8 picomolar (pM) as suggested in De Fanti et al. (2016).

4.2.10. Library amplification

Libraries with low concentrations were re-amplified using the Platinum PCR SuperMix High Fidelity included in the Ion Plus Fragment Library Kit (Life Technologies, Carlsbad, CA, USA). First, 5 µL of Low TE buffer was added to the ~20 µL of the purified adapter-ligated library, and then 130 µL of reaction mix (Table 4.2.10.1) was prepared and split into two tubes.

Reagent	*Volume for 1 reaction (µl)*
Platinum® PCR SuperMix High Fidelity	100
Library Amplification Primer Mix	5
unamplified library	25
Total	130

Table 4.2.10.1 | Library amplification.

The temperature profile of the thermal cycler program was as follows: initial denaturation step at 95° C for 5 minutes, 7 cycles consisting of 15 seconds at 95° C, 15 seconds at 58° C, 1 minute at 70° C, followed by a final holding for up to 1 hour at 4° C.

Since it is in general preferable not to amplify a library, as to avoid the introduction of PCR-induced errors, in this step is important to minimize the numbers of cycles. 7 is the number of cycles suggested by the manufacturer instructions for 20 ng of input unamplified library. Finally, the previously split PCRs of the same sample were combined in a new 1.5-mL tube. To remove the residuals of the reagents of the previous reaction, the library was purified using 195 µL Agencourt® AMPure® XP Reagent (Beckman Coulter, Brea, CA, USA) in the first step of the process. The amplified libraries concentrations were determined by qPCR as previously described.

4.2.11. Emulsion PCR

The Ion PGM™ Hi-Q™ 0T2 Kit (Life Technologies, Carlsbad, CA, USA) was used on a OneTouch™ 2 instrument (Life Technologies, Carlsbad, CA, USA) to prepare template-positive Ion Sphere™ Particles (hereafter ISPs) containing clonally amplified DNA.

Emulsion PCR uses a water-in-oil emulsion, where the water acts as a dispersed phase and the oil as a dispersion medium. This kind of emulsion is obtained by combining a determined volume of water with a larger volume of oil. Each water droplet acts as a microreactor. This emulsion happens in the reaction filter and, in an ideal situation, in each microreactor, there

will be exactly one magnetic bead coated with primers that are complementary to the adapter P1, one strand of library DNA, primers and PCR mix. Inside each droplet, a PCR reaction occurs: first, a denaturation step of the library fragment; then, an annealing step where the reverse strand anneals to adapter site on the beads; after, an extension phase, when the polymerase amplifies the forward strand starting from the beads towards the primer site. This PCR reaction takes place in the vertical thermal cycler integrated into the Ion OneTouch™ 2 Instrument, inside the amplification plate where the droplets are moved, and this system enables thermal cycling of the microreactors. At the end of the process, the ISPs contain thousands of copies of the single library fragment originally present and all the ISPs are recovered and precipitated through an integrated centrifuge.

In detail, the library dilutions were prepared according to the dilution factors calculated from the qPCR results and then 5 µL of each diluted library were pooled together. The samples are pooled together in equimolar quantities to maximize the chance that all will be sequenced with the same coverage. Afterwards, the Ion PGM™ Hi-Q™ ISPs (Life Technologies, Carlsbad, CA, USA) were centrifuged and pipetted up and down to make sure they were thoroughly mixed. Immediately after the amplification solution was prepared (Table 4.2.11.1), adding all reagents to a 2-mL tube already containing 800 µL of Ion PGM™ Hi-Q™ Reagent Mix.

Order	*Reagent*	*Volume (µL)*
1	Nuclease-free water	25
2	Ion PGM™ Hi-Q™ Enzyme Mix	50
3	Diluted library (not stock library)	25
4	Ion PGM™ Hi-Q™ ISPs	100
	Total	1000

Table 4.2.11.1 | Emulsion PVR reaction mix.

Considering that the run of the Ion OneTouch™ 2 instrument must start less than 15 minutes after preparing the amplification solution, it is really important to have already initialized and cleaned the instruments and refilled the reagents when this step of the process is reached.

With the Ion OneTouch™ 2 instrument ready the Ion OneTouch™ Reaction Filter was filled with 1000 µL of the amplification solution and 1.6 mL of Ion OneTouch™ Reaction Oil. To add the reagents, the filter was inverted upside down, to be loaded it was carefully

inverted, to avoid an emulsion of the two phases at this stage, a situation that may cause some troubles during the experiment. Finally, the filter was inserted into the instrument and the run was set and started. As above mentioned, after the amplification reaction was over, the sample was spun for 10 minutes, to form a ISPs pellet inside two tubes containing a recovery solution.

4.2.12. Enrichment

The enrichment of clonally amplified libraries was achieved with Ion PGM™ Enrichment Beads (Life Technologies, Carlsbad, CA, USA) on an Ion OneTouch™ ES instrument (Life Technologies, Carlsbad, CA, USA).

In detail, most of the supernatant from the latter centrifugation was removed and the ISPs were resuspended in the remaining Ion PGM™ OT2 Recovery Solution (Life Technologies, Carlsbad, CA, USA). Then, 500 µL of Ion OneTouch™ Wash Solution (Life Technologies, Carlsbad, CA, USA) was added to each Recovery Tube, and the two aliquots of the emulsion PCR product were combined and centrifuged at 15,500 g for 2.5 minutes. The majority of the supernatant was once again discarded. The reagents used to fill the 8-well strip of the OneTouch™ ES instrument (Life Technologies, Carlsbad, CA, USA) were a Melt-Off solution (Table 4.2.12.1), all but 4 µL of this unenriched ISPs solution, that are retained for quality assessment and the washed and resuspended Dynabeads® MyOne™ Streptavidin C1 Beads.

Order	*Reagent*	*Volume (µL)*	*Final composition*
1	Tween® Solution	280	0.1% Tween® detergent
2	1 M NaOH	40	125 mM NaOH
	Total	320	

Table 4.2.12.1 | Melt-Off solution.

The 8-well strip was then filled (Table 4.2.12.2) and inserted in the instrument with the square-shaped tab on the left and was pushed all the way to the right end of the slot of the tray.

Well number	*Reagent to dispense in well*
Well 1 (closest to square-shaped tab)	Template-positive ISP sample
Well 2	130 µL of DynabeadsR MyOne™ Streptavidin C1 Beads resuspended in MyOne™ Beads Wash Solution
Well 3	300 µL of Ion OneTouch™ Wash Solution
Well 4	300 µL of Ion OneTouch™ Wash Solution
Well 5	300 µL of Ion OneTouch™ Wash Solution
Well 6	Empty
Well 7	300 µL of freshly-prepared Melt-Off Solution
Well 8	Empty

Table 4.2.12.2 | OneTouch™ ES strip.

4.2.13. Quality Control assay

The quality of the recovered template Ion Sphere Particles (ISPs) was evaluated using the Ion Sphere Quality Control Kit on the Qubit® 2.0 Fluorometer (Invitrogen™ Life Technologies - Carlsbad, CA, USA) by means of two fluorophores: Alexa Fluor® 488 and Alexa Fluor® 647 (Thermo Fisher Scientific Company, Waltham, USA).

In detail, a probe labelled with Alexa Fluor® 488 anneals to primer B sites, or all of the ISPs present, while a probe labelled with Alexa Fluor® 647 anneals to primer A sites, or only the ISPs with extended templates. The ratio of the Alexa Fluor® 647 fluorescence (templated ISPs) to the Alexa Fluor® 488 fluorescence (all ISPs present) yields the % templated ISPs. The RFU values obtained by the calibration of Alexa Fluor® 488 and Alexa Fluor® 647 Calibration Standard reagents, once reported in the Qubit® Easy Calculator, permit the calculation of the Calibration Factor specific for the Qubit® 2.0 Fluorometer. To measure the template unenriched sample, Ion OneTouch™ Wash Solution from the Ion PGM™ Hi-Q™ OT2 Solutions were added to the unenriched Ion PGM™ Hi-Q™ Ion Sphere™ Particles, in order to reach a final volume of 100 µL. Afterwards, 2 µL of the solution was transferred into a 0.2-mL PCR tube to which were added 19 µL of annealing buffer and 1 µL of Ion Probes. Then, the tubes were loaded into a thermal cycler, the following protocol was performed to anneal the Ion Probes: 95° C for 2 minutes and 37° C for 2 minutes.

The unbound probes were then removed by three washes performed by adding 200 µL of Quality Control Wash Buffer, centrifuging at 15,500 g for 1.5 minutes and removing all the supernatant except for 10 µL. After the final wash, 190 µL of Quality Control Wash Buffer was added to the sample. The sample was measured by its AF 488 value and its AF 647 value. These were reported in the Qubit® Easy Calculator Microsoft® Excel® Spreadsheet file containing the Calibration Factor specifically calculated for the Qubit® 2.0 Fluorometer used

and permitted a templated ISP evaluation. In general, as a first indication, an RFU Alexa FluorR 488 above 100 counts is and an acceptable value, under 100 counts usually means that in the assay there are no or very few ISPs.

4.2.14. High-throughput sequencing

Sequencing reaction of the recovered templates was performed on an Ion Torrent PGMTM System (Life Technologies, Carlsbad, CA, USA) using the Ion PGM™ Hi Q™ Sequencing Kit (Life Technologies, Carlsbad, CA, USA) and the Ion 316™ Chip v2 (Life Technologies, Carlsbad, CA, USA).

Since the Ion Personal Genome Machine™ (PGM™) sequencer technology is based on pH variation, before starting the experiment, the instrument was first cleaned with chlorite and then with 18MΩ water. After that, an initialization step was performed. The instrument uses 350 µL of freshly prepared 100 mM NaOH (store in bottle W1) to match bottle W2 pH solution (prepared by diluting 70 µL of a 100 mM NaOH solution and the entire bottle of Ion PGM™ Hi-Q™ Sequencing W2 Solution in ~2 L of 18 MΩ water) to the known pH value of ~50 mL of the Ion PGM™ Sequencing W3 Solution (stored in bottle W3), which has a known pH.

Since pH variation is critical in this process, the instrument must operate in an Argon environment, paying attention to preventing atmospheric CO_2 from reducing the pH of Wash Solution 2 Bottle solution.

The dNTPs solutions were prepared by transferring 20 µL of each dNTP stock solution into its respective 50 mL Reagent Bottle. Then the instrument filled each Reagent Bottle with 40 mL of W2 Solution. For the sample preparation, 5 µL of Control ISPs were directly added to the entire volume of enriched template-positive ISPs. These Control ISPs contain test fragments, that is to say, known sequence strand that permits to test the efficiency of the run and the reliability of the called bases. Then the sample was thoroughly mixed, centrifuged at 15,500g for 2 minutes and the supernatant was removed leaving in the 0.2-mL tube just ~15 µL of the sample. Afterwards, 12 µL of Sequencing Primer were added to the ISPs and the sample was placed in the thermal cycler to let the sequencing primer annealing to the templates. The thermal cycler program was set as follows: 95° C for 2 minutes and 37° C for 2 minutes.

After testing the chip to ensure that it was functioning properly prior to loading the sample, 3 µL of Ion PGM™ Hi-Q™ Sequencing Polymerase was added to the ISPs, and the sample

was incubated at room temperature for 5 minutes. At the end of this step, the sample has now a total volume of 30 µL. The retained liquid from the check procedure needs to be completely discarded, and this was accomplished through pipetting and several centrifuging steps.

To load the Ion 316™ Chip the entire volume of the sample was required: by means of the Rainin® SR-L200F pipette the sample was slowly loaded into the Chip (~ 1 µL per second). Then the chip was centrifuged for 30 seconds with the chip tab pointing in and 30 seconds with the chip tab pointing out.

This slow loading process and the two centrifugation steps ensure that the sample solution was added uniformly to the surface of the chip, to increase the possibility that all the wells were filled with an ISP. As all the ISPs were expected at the bottom of each well, the sample solution was pipetted out of the chip. The remained liquid was further discarded with a chip centrifugation step upside-down. Finally, the chip was loaded into the instrument and the sequencing run was started.

4.2.15. Sanger sequencing

For the subset of samples tested twice (see mtDNA amplification), Sanger sequencing experiment was performed on a 3730 DNA Analyzer (Applied BioSystems, Foster City, USA) at the "Unità Operativa di Genetica Medica dell'Azienda Ospedaliera di Bologna", to verify the repeatability of the mitochondrial data and to confirm the authenticity of the NGS sequencing results.

In detail, the purified PCR product was added to the sequencing reaction mixture (Table 4.2.15.1), with a volume ranging from 0.5 to 1.5 µL, determined based on the intensity of the bands on the electrophoresis gel. For the sequencing reaction the thermal cycler was as follows: 30 cycles consisting of 2 minutes and 10 seconds at 96°C, 15 seconds at 50°C and 4 minutes at 60°C.

Reagents	*Stock concentration*	*Final concentration*	*Volume for 1 sample (µL)*
Nuclease-free water			3,9/ 3,4/ 2,9 a volume
Big Dye Terminator (v1.1)			1
Buffer	5X	2.5X	3
Primer Fwd	10 µM	1 µM	1,6
Primer Rvs	10 µM	1 µM	1,6
Purified DNA			0,5/ 1/ 1,5 depending on PCR quality results
Total			10

Table 4.2.15.1 | Sequencing reaction mixture.

Then, the reaction products were purified by the means of a precipitation mixture (Table 4.2.15.2) that, combined with a long centrifugation at 3,000 g at 4° C for 30 minutes, let the DNA precipitate at the bottom of the well plates. The supernatant was discarded by quickly inverting the plate over the sink and another wash, this time just with 70 µL of ethanol 70% was carried out. Several centrifugations followed, even with the plate upside down on an absorbent pad, to remove all supernatant and guarantee no presence of residual ethanol. Before loading the plate into the sequencer, 20 µL of injection solution was added to each well to resuspend the sample and a silicone septa mat was used to cover the plate.

Reagent	*Volume for 1 sample (µL)*
Nuclease-free water	10
Sodium acetate	2
Ethanol 100 %	55
Total	67

Table 4.2.15.2 | Precipitation reaction.

4.2.16. Sequence analysis

The chromatograms obtained by Sanger sequencing were manually checked with Chromas 2.4.4. (Technelysium) and then a preliminary analysis of the mutations compared to the rCRS (Anderson et al., 1981; Andrews et al., 1999) was performed with DNA Alignment (fluxus-engineering.com; http://www.fluxus-engineering.com/align.htm).

Regarding the NGS sequencing, the Ion Torrent data for each barcoded library were processed using a customized pipeline constituted by tools implemented in the Galaxy 16.01 platform (https://usegalaxy.org/) (Giardine et al., 2005; Blankenberg et al., 2010; Goecks et al., 2010).

The FASTQ files obtained for each sample were filtered for quality score and length using the following parameters set: minimum size (nt = 168), minimum quality (Q = 20.0) and a maximum number of bases allowed outside of quality range (n = 5). The filtered FASTQ was later converted into a FASTA file and then the three overlapping fragments were separated by the means of the "Barcode Splitter" tool of the platform. The primer sequences were indicated as barcodes specifying that they were at the 5' end of the sequences, up to 5 mismatches and up to 5 barcodes nucleotide deletions were allowed. All the forward and reverse sequences of each fragment were unified paying attention to reverse-complement the reverse sequences and subsequently to remove, by the means of the "clip" tool the primer sequences. Sequences

shorter than 100 bp were discarded, as well as the non-clipped reads. The "collapse" tool was used to collapse all the identical sequences into a single sequence, carrying in the output the information of the count for each haplotype. At the end of this filtering pipeline, those haplotypes with total counts lower than 5 were removed.

4.2.17. SNPs genotyping

A total of 22 SNPs (Table 4.2.17.1a; 4.2.17.1b) in the mtDNA coding region were selected to confirm the haplogroup assignment preliminarily inferred with the HVS-I haplotype motifs. Coding region SNPs were genotyped by means of two different multiplex PCR (Bertoncini et al., 2011): *multiplex 1* included variants that define the most common non-H European lineages (4216L, 4529L, 4580L, 7028L, 10398L, 10400H, 10873H, 12308L, 12705L, 14766L) (Richards et al., 2000), whereas *multiplex 2* contained variants of H sub-lineages (3010L, 3915H, 3936H, 3992L, 4310L, 4745L, 4336L, 4769H, 4793H, 6776H, 13708L, 13759L) (Herrnstadt et al., 2002).

	Name	*SNP primer sequences 5' - 3'*	*SNP lenght*	*SNP fc (µM)*
Multiplex 1	*4216*	CTCTACACAACATATTTTGTCACCAAG GGTTTGAGGGGGAATGCTGGAG	195	0.3
	4529-4580	CAACCCGTCATCTACTCTACCAT CTTCTGTGGAACGAGGGTTTATT	148	0.3
	7028	CACCGTAGGTGGCCTGACTGGC GTGTAGCCTGAGAATAGGGG	168	0.15
	10398-10400	AAATTGCCCTCCTTTTACCCCTA TGTAAATGAGGGGCATTTGG	224	0.4
	10873	CATAATTTGAATCAACACAACCACC GTTAGGGGGTCGGAGGAAAAGGTTG	123	0.1
	12308	CTGCTAACTCATGCCCCCATG ATTACTTTTATTTGGAGTTGCACCAAGATT	106	0.4
	12705	TGTAGCATTGTTCGTTACATGG AGTTGGAATAGGTTGTTAGCGG	147	0.2
	14766	TCAACTACAAGAACACCAATGACC GGAGGTCGATGATGAGTGG	82	0.4
Multiplex 2	*3010*	CAATAACTTGACCAACGGAACA CGGTCTGAACTCAGATCACGTA	180	0.15
	3915-3992c	TAGCAGAGACCAACCGAACC GAAGATTGTAGTGGTGAGGGTGT	158	0.15
	4336	GGAGCTTAAACCCCCTTATTTC GATAGGTGGCACGGAGAATTT	80	0.15
	4769-4793c	CCGGACAATGAACCATAACC TGGGTAACCTCTGGGACTCA	118	0.15
	6776	GCTTCCTAGGGTTTATCGTGTG	140	0.15
	13708-13759	AACGAAAATAACCCCACCCTA GTTGTTTGGAAGGGGGATG	113	0.20

Table 4.2.17.1a | PCR primer pairs used for amplification of SNPs in coding region of the mtDNA.

Name	*SBE primer sequences 5' - 3'*	*SBE mutation*	*SBE lenght*	*Minisequencing SBE primer fc (µM)*
3010	CTCGATGTTGGATCAGGACATCCC	G–A	24	0.1
3915	TGACTGACTAAGCCTGAGACTAGTTCGGACTC	G–A	32	0.2
3992	TGACTGACTGACTGACTCCCTATTCTTCATAGCCGAATACA	C–T	45	0.3
4216	CCCCTACCACTCACCCTAGCATTACTTATATGA	T–C	33	0.2
4336	CTGACTGACTGACTGACTGACTGACTGACTGCTTAAACCCCCTTATTTCTAGGAC	T–C	55	0.2
4529	CTTTGCAGGCACACTCATCAC	A–T	21	0.15
4580	TTACCTGAGTAGGCCTAGAAATAAACAT	G–A	28	0.1
4769	GACTGACTGACTGGGCTATTCCTAGTTTTATTGCTATAGC	A–G	40	0.3
4793	GACTGACTGACTGACTGACTGACTGACTGAACTCAGAAGTGAAAGGGGGC	A–C	50	0.4
6776	ACTGACTGACTGACTGACTGACTCGTGTGTCTACGTCTATTCCTACTGTAAATAT	T–C	55	0.3
7028	TACACGACACGTACTACGTTGTAGC	C–T	25	0.1
10398	GACTGACTGACTGACTGACTGACTGACTATGAGTGACTACAAAAAGGATTAGACTGA	A–G	57	0.5
10400	CCCCCCCCCCCGTTTTGTTTAAACTATATACCAATTC	C–T	37	0.5
10873	CCCCCCCCCCCCCCCGTTGTTGTTGATTTGGTTAAAAAATAGTAG	T–C	45	0.1
12308	CCCCCCCCCCCCCCCCCCCCCCCATTGGTCTTAGGCCCCAA	A–G	41	0.4
12705	CCCCCCCCCCCCCCCCCAACATTAATCAGTTCTTCAAATATCTACTCAT	C–T	49	0.3
14766	AATGACCCCAATACGCAAAA	C–T	20	0.3
13708	(gact)13 gaCTACTAAACCCCATTAAAGGCCTG	G-A	78	0.03
13759	(gact)10TTCTCATTACTAACAACATTTCCCCC	G-A	66	0.01
3936	(gact)14gTGCGGCGTATTCGATGTTGAA	C-T	78	0.06
4310	TCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGG	A-G	38	0.04
4745	gactAATGAACCATAACCAATACTACCAATCA	A-G	32	0.02

Table 4.2.17.1b | PCR primer pairs used for amplification of SNPs in coding region of the mtDNA.

PCR was performed in a volume of 5 µL using 2.5 µL of PCR Master Mix (Qiagen GmbH, Hilden, Germany), 0,25 µL of 20X *multiplex 1*/*multiplex 2* primers, 1 µL of DNA template and 1,25 µL of nuclease-free water (Qiagen GmbH, Hilden, Germany). Cycle conditions were: 95° C for 15 minutes, 40 cycles of 94° C for 30 seconds, 58° C for 60 seconds, 72° C for 60 seconds, and a final step of 72° C for 10 minutes.

PCR products were checked on 1.5% agarose gel and subsequently 1.5 µL of PCR products from both the multiplex-PCR were purified with 1.5 µL of ExoSAP-IT™ PCR Product Cleanup Reagent (Applied BioSystems, Foster City, USA) with the following cycle conditions: incubation step at 37°C for 15 minutes and inactivation step by heating at 85° C for 15 minutes.

Afterwards, a single-base extension (SBE) assay was performed in a 5 µL volume using 2 µL of SNaPshot ready reaction mix (Applied BioSystems, Foster City, USA), 0.25 µL of *multiplex 1* and *multiplex 2* primers mix, 0.5 µL of nuclease-free water (Qiagen GmbH, Hilden, Germany) and 2 µL of purified amplicons from both multiplex PCR. SBE amplification condition consisted of 25 cycles for 10 seconds at 96°C, 5 seconds at 50°C and 30 seconds at 60°C.

The obtained product was purified adding 1 µL of shrimp alkaline phosphatase enzyme (0.2 Units/mL) to each reaction volume and then the solution was heated with the following temperature condition: dNTPs and primers degradation at 37°C for 60 minutes, SAP

inactivation at 80°C for 15 minutes. Capillary electrophoresis reaction was performed at the Department of Diagnostic and Laboratory Services and Legal Medicine (University of Modena and Reggio Emilia) on an ABI PRISM™ 3130 DNA Genetic Analyzer (Applied BioSystems, Foster City, USA), and the obtained data were analysed using GeneScan 3.7 software (Applied BioSystems, Foster City, USA).

4.2.18. Autosomal analysis

A potential kinship relationship between two couples of individuals buried in the same grave (grave 155: individuals NOR10a and NOR10b; grave 171: individuals NOR3a and NOR3b) was investigated using two commercial forensic PCR kits: the Globalfiler™ assay kit (Thermo Fisher Scientific Company, Waltham, USA) and the DIPplex® kit (Qiagen GmbH, Hilden, Germany). The autosomal analyses were performed at the Institute of Forensic Sciences "Luis Concheiro", University of Santiago de Compostela, according to the manufacturer's protocol for the treatment of degraded and problematic DNA samples.

The GlobalFiler™ kit (Thermo Fisher Scientific Company, Waltham, USA) which incorporates 21 autosomal short tandem repeat (STR) loci, namely CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, VWA, D2S1338, D19S433, D1S1656, D12S391, D2S441, D10S1248, TPOX, D22S1045, SE33 and additional three gender determination loci (Amelogenin, Yindel, and DYS391) was amplified as suggested in scientific reference literature (Hennessy et al., 2014; Wang et al., 2015). For each sample, the analyses were conducted on two different extracts, for each of which, an independent amplification was performed.

PCR products were separated and detected by capillary electrophoresis in an ABI 3500 Genetic Analyzer (Thermo Fisher Scientific Company, Waltham, USA), and genotyping was done using allelic ladders provided with the kit and the GeneMapper® ID-X software v1.4 (Thermo Fisher Scientific Company, Waltham, USA). Allele nomenclature, quality control and statistical issues follow the recommendations of the International Society for Forensic Genetics (Barr et al., 1997; Lincoln et al., 1997). GlobalFiler STR loci comprise all the core loci of the European Standard Set (ESS), and all the 20 autosomal loci required in the expanded Combined DNA Index System (CODIS) core STR loci (Hammond et al., 1994; Butler et al., 2012).

To overcome the STRs typing limitation, alternative approaches based on single-nucleotide polymorphisms (SNPs) and short deletion/insertion polymorphisms (InDels) have been

developed, establishing a powerful supporting typing system for forensic science (Pereira et al, 2009). Thirty autosomal insertion-deletion (InDel) polymorphisms and the amelogenin locus were analysed for the same samples as the autosomal STRs analysis, using the Investigator® DIPplex Kit (Qiagen GmbH, Hilden, Germany), following the same workflow as STR assays (LaRue et al., 2012). The PCR amplification was performed in a GeneAmp PCR system 9700 thermal cycler (Life Technologies Carlsbad, CA, USA), and the InDels typing were performed via Applied Biosystems 3130 Genetic Analyzer® following the manufacturer's recommendations. Allele allocation was carried out with GeneMapper ID 3.2.1 analysis software (Life Technologies Carlsbad, CA, USA) using the allelic ladder and the set of bins and panels provided by the kit, and the DIPSorter freeware (Qiagen GmbH, Hilden, Germany) was used for an easy and accurate interpretation of results.

4.3. Statistical analyses

4.3.1. Intra-population analysis

4.3.1.1. Haplogroup assignment

All merged and quality-filtered sequences of Novilara samples were edited and aligned to the revised Cambridge Reference Sequence (rCRS, GenBank Accession Number NC 012920) (Anderson et al., 1981; Andrews et al., 1999) with DNA Alignment (fluxus-engineering.com), BioEdit v7.2.5 (Hall, 1999) and MEGA7 (Kumar et al., 2016) software. Afterwards, the types and frequency of nucleotide variations were checked among the collapsed reads for every single individual, such as C->T transitions, which represent the prevalent signal of post-mortem miscoding lesions in authentic aDNA (Stiller et al., 2009; Bollongino et al., 2013; Dabney at al., 2013b). Mitochondrial haplogroups were determined based on the PhyloTree mtDNA phylogeny, built 17 (www.phylotree.org) (van Oven and Kayser, 2009) and Haplogrep2 software (Kloss-Brandstätter et al., 2011).

4.3.1.2. Summary and population differentiation statistics

To evaluate the intra-population genetic variability, summary statistics analyses were computed with the Arlequin software ver. 3.5 (Berne, Switzerland) (Excoffier and Lischer, 2010). The following parameters were studied:

- Number of different haplotypes (k): number of different haplotypes present in the population.
- Gene diversity (H): it is defined as the probability that two randomly chosen haplotypes are different in the sample and it is equivalent to the expected heterozygosity for diploid data. Gene diversity is computed as

$$\hat{H} = \frac{n}{n-1}(1 - \sum_{i=1}^{k} p_i^2)$$

where n is the number of sequence copies in the sample, k is the number different haplotypes present and p_i is the sample frequency of the i-th haplotype (Nei, 1987).

- Mean number of pairwise differences (π): it is the mean number of differences between all pairs of haplotypes in the sample. It is calculated as

$$\hat{\pi} = \frac{n}{n-1} \sum_{i=1}^{k} \sum_{j=1}^{k} p_i p_j \hat{d}_{ij},$$

where d_{ij} is an estimate of the number of mutations having occurred since the divergence of haplotypes i and j, k is the number of haplotypes, and p_i is the frequency of haplotype i, and n is the sample size (Tajima, 1983; 1993).

- Nucleotide diversity (π_n): describes the probability that two randomly chosen homologous nucleotides are different from one another. It is equivalent to the gene diversity at the nucleotide level and it is calculated as the mean number of pairwise differences over the number of analysed loci

$$\hat{\pi}_n = \frac{\sum_{i=1}^{k} \sum_{j<i} p_i p_j \hat{d}_{ij}}{L}$$

where d_{ij} is an estimate of the number of mutations having occurred since the divergence of haplotypes i and j, k is the number of haplotypes, p_i is the frequency of haplotype i and L is the number of loci analysed (Tajima, 1983; Nei, 1987).

4.3.2. Inter-population variability

4.3.2.1. Populations used in comparative analyses

To evaluate the inter-population genetic variability, HVS-I data (nps 16024-16383) obtained from the ancient Piceni from Novilara were compared to 12 European prehistoric populations available from literature (populations ranging from the Bronze Age (n = 6) to the

Iron Age (*n* = 6) period) (Table 4.3.2.1.1), as well as to 1833 present-day unrelated individuals from continental Italy, Sicily, and Sardinia, clustered into 8 macro-areas following the strategy approach described in Boattini et al. (2013) (Table 4.3.2.1.2).

Abbreviation	*Full name*	*Locality*	*Date Range*	*n*	*Reference*
CRE_BA	Minoan	Crete, Greece	4,400-3,700 yBP	37	Hughay et al., 2013
NUR_BA	Nuragic	Sardinia, Italy	4,300-3,000 yBP	39	Caramelli et al., 2007; Der Sarkissian, 2011
SPA_BA	Catalan	Spain	3,200 BC	7	Simòn et al., 2011
TK_BA	Titriş Höyuök individuals	Turkey	2,300-2100 yBP	12	Matney et al., 2012
CAT_BA	Catacomb culture	North Pontic Region	4,700-4,000 yBP	24	Wilde et al., 2014
YAM_BA	Yamnaya	North Pontic Region	5,000-4,500 yBP	25	Wilde et al., 2014
SCYR_IA	Scythians	Russia	2,200-2,600 yBP	16	Der Sarkissian, 2011
SCYU_IA	Scythians	Moldova/Ukraine	4th-2nd BC	19	Juras et al., 2017
GER_IA	La Tène	Germany	6th-4th century BC	10	Knipper et al., 2014
PRZ_IA	Przeworsk and Wielbark culture	Poland	200 BC-500 AD	23	Juras et al., 2014
PIC_IA	**Piceni**	**Italy**	**8th-7th c BC**	**9**	**present study**
ETR_IA	Etruscan	Italy	7th-3rd c. BC	23	Vernesi et al., 2004; Ghirotto et al., 2013
SPA_IA	Spanish	Girona	8tht-6th c BC	17	Sampietro et al., 2005

Table 4.3.2.1.1 | mtDNA reference dataset for the 12 European prehistoric populations included into MDS analyses along with the data obtained in this study. Abbreviation: BA, Bronze Age; IA, Iron Age.

Population	*Area*	*n*	*Reference*
Cuneo	1	40	Boattini et al., 2013
Como	1	39	Boattini et al., 2013
Brescia	1	40	Boattini et al., 2013
Savona/Genova	1	43	Boattini et al., 2013
Aviano	2	29	Boattini et al., 2013
Vicenza	2	40	Boattini et al., 2013
Treviso	2	39	Boattini et al., 2013
Udine	2	51	Brisighelli et al., 2012
Bologna	3	100	Bini et al., 2003
Modena	3	43	Turchi et al., 2007
Grosseto/Siena	4	36	Boattini et al., 2013
Pistoia	4	16	Boattini et al., 2013
Casentino	4	120	Achilli et al., 2007
Murlo	4	86	Achilli et al., 2007
Volterra	4	113	Achilli et al., 2007
Terni	4	31	Boattini et al., 2013
Macerata	5	39	Boattini et al., 2013
Foligno	5	40	Boattini et al., 2013
Ascoli Piceno	5	53	Brisighelli et al., 2012
Piceni from Novilara	**5**	**9**	**Present study**
Ancona	5	70	Babalini et al., 2005
L'Aquila	6	25	Boattini et al., 2013

Table 4.3.2.1.2 | mtDNA reference dataset for the 34 current Italian populations included into comparison analyses along with the data obtained in this study.

Population	Area	n	Reference
Benevento	6	36	Boattini et al., 2013
Cosenza/Catanzaro/Crotone	6	37	Boattini et al., 2013
Belvedere	6	50	Brisighelli et al., 2012
Campobasso	6	37	Boattini et al., 2013
Matera	6	36	Boattini et al., 2013
Lecce	6	39	Boattini et al., 2013
Sanniti	6	50	Brisighelli et al., 2012
Lucera	6	60	Brisighelli et al., 2012
Agrigento	7	42	Boattini et al., 2013
Catania	7	37	Boattini et al., 2013
Ragusa/Siracusa	7	39	Boattini et al., 2013
Trapani	7	40	Brisighelli et al., 2012
Sardinia	8	233	Richards et al., 2000; Falchi et al., 2006; Di Rienzo & Wilson, 1991
Italian macro-area			
IT_Area 1	-	162	-
IT_Area 2	-	163	-
IT_Area 3	-	143	-
IT_Area 4	-	371	-
IT_Area 5	-	233	-
IT_Area 6	-	370	-
IT_Area 7	-	158	-
IT_Area 8	-	233	-

Table 4.3.2.1.2 | mtDNA reference dataset for the 34 current Italian populations included into comparison analyses along with the data obtained in this study.

4.3.2.2. Genetic Distances

The measures of genetic distance are statistics that allow inferring the evolutionary relationships between populations or molecules. A commonly used classical measure of genetic distance between populations is the F_{ST} methods that can be applied to all kind of data. It varies between 0, for identical populations, and 1 for populations that share no alleles. It is based on allelic frequencies, it is calculated between pairs of populations and represents the excess of homozygotes in the subpopulations with respect to the metapopulation. It is calculated as

$$Fst = Vp / p (1-p)$$

where Vp is the variance of the frequency of the i allele in the metapopulation, and p is the mean frequency of the i allele between the populations.

Pairwise Fst distances were computed on HVS-I data (nps 16024-16383) using Kimura's two-parameter distance option (Kimura, 1980) as implemented in Arlequin software ver. 3.5

(Berne, Switzerland) (Excoffier and Lischer, 2010). It outputs a corrected percentage of nucleotides for which two haplotypes are different. The correction also allows for multiple substitutions per site but takes into account different substitution rates between transitions and transversions. The transition-transversion ratio is estimated from the data (Kimura, 1980; Jin and Nei, 1990)

In the construction of the matrix, a value for the shape parameter a of the gamma function is also set, when selecting a distance allowing for unequal mutation rates among sites. For the study of mitochondrial DNA, the a parameter was set on 0,26.

4.3.2.3. Multidimensional scaling (MDS)

The Multidimensional Scaling (Kruskal, 1964) is a mathematical procedure that allows the representation of the objects under study in a Euclidean space, defined by a desired number of dimensions, so that the distances reproduced reflect the values observed in the best way possible. The method proceeds through a series of iterations moving around objects in the space defined by the requested number of dimensions and checking how well the distances between objects can be reproduced by the new configuration. The goal is to maximize the goodness-of-fit, which is represented by the stress value, defined as follows:

$$\text{Phi} = \sum (d_{ij} - \delta_{ij})2$$

where d_{ij} represents the observed distance between objects, while δ_{ij} is the reproduced distance. Higher the similarity between the two matrices, the observed and the reconstructed one, lower will be the value of stress. The iterations stop when this parameter cannot decrease further, reaching a threshold value. Slatkin Fst-values (Slatkin, 1995) were used to reconstruct non-metric Multidimensional scaling plot visualized in bi-dimensional space using R 'MASS' package (R-DevelopmentCoreTeam 2008).

4.4. Results and Discussion

4.4.1. Next Generation Sequencing

- *Chip loading*: the NGS run provided 311 Mb. Figure 4.4.1.1 shows a heat map concerning the chip loading. 63% of the wells were loaded with a ISP. The legend

shows the percentage of loading, from blue indicating 0% to red corresponding to 100%.

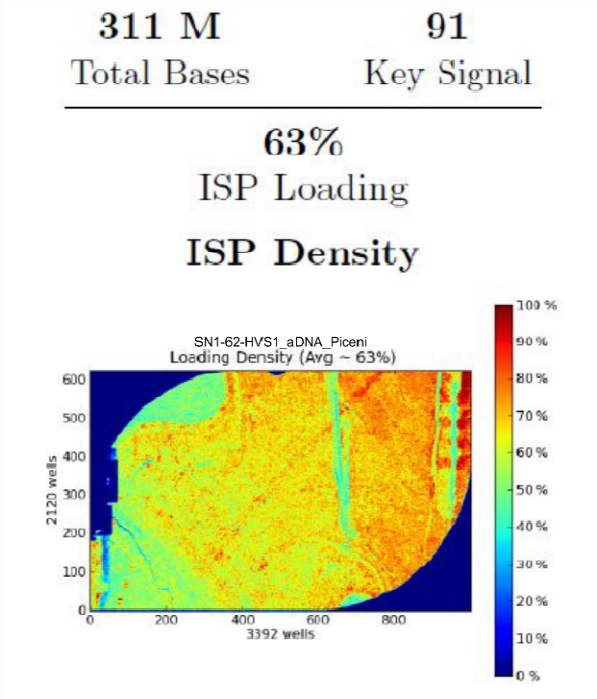


Figure 4.4.1.1 | Chip loading map.

- *Run and alignment summary*: a total of 2,174,260 reads were produced and the 55% of these were usable for downstream analyses (Figure 4.4.1.2a). Focusing on the ISP, as previously said, just 37% of the chip's wells were empty and all the ISPs were enriched (i.e. none of the ISPs contained in our sample did not carry any template). 62% of the ISPs contain only one copy of DNA template and thus, 38% were polyclonal. Finally, the last data shown in the figure is about the nature of the reads: 87% constituted the final library produced by the experiment, while the remaining 13% either carried test fragments, contained adapter dimer or were of low quality. As shown in Figure 4.4.1.2b, 276 of the 311 Mb produced aligned to the reference that was indicated to the Torrent Server (human genome hg_19), resulting in an average coverage of 0.1X. This last data means that, on average, each base of the human genome has been sequenced 0.1 times. Summarizing, of the 2,005,960 bases produced by the 62% monoclonal ISPs, 96% aligned to the reference, specifically and very efficiently to the target region in the Chromosome M that was specified to the Torrent Server with a BED file of the HVS-I.

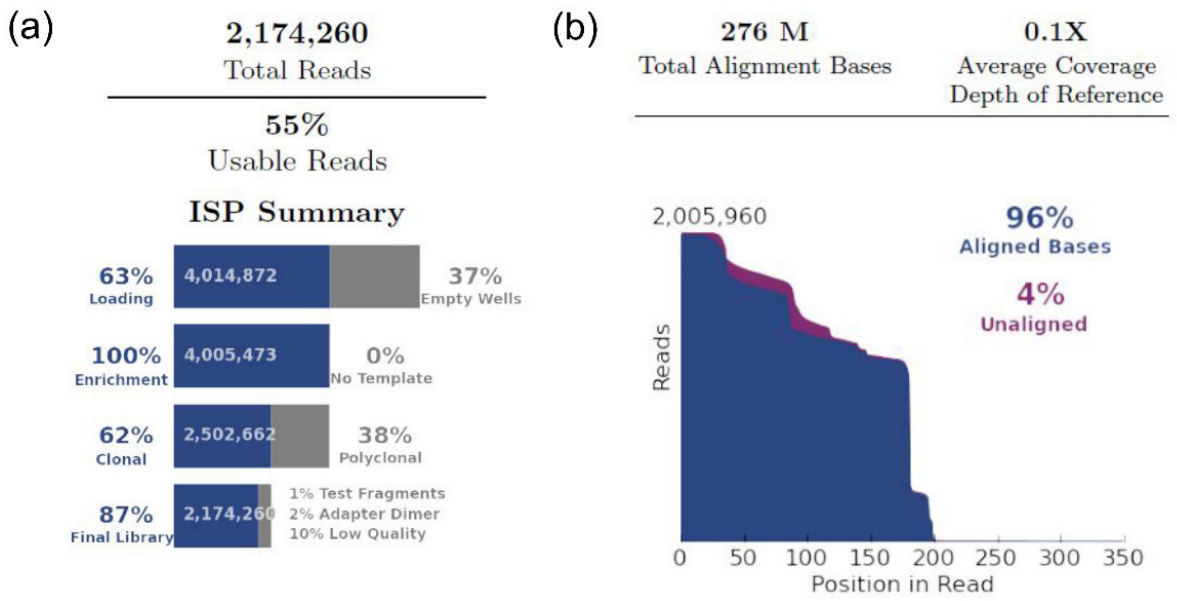


Figure 4.4.1.2 | (a) Run summary, (b) Alignment summary.

- *Ion Torrent sequencing*: The Torrent Server produced a "Coverage Analysis Report" for each sample that provided summary information about the number of mapped reads, the percent of reads on target, the average base coverage depth, and the uniformity coverage (Table 4.4.1.3).

Sample	*Bases*	*Reads*	*Mean Read Length*	*no. of mapped reads*	*Percent reads on target*	*Average base coverage depth*	*Uniformity of base coverage*
NOR3a	8,538,360	57,248	149bp	56,942	99.81%	18,737	86.00%
NOR3b	13,866,513	91,125	152 bp	90,733	99.82%	30,640	86.00%
NOR8	10,216,388	76,465	134 bp	76,036	99.41%	22,534	100.00%
NOR9	8,628,478	60,426	143 bp	59,981	97.78%	18,114	100.00%
NOR10a	12,494,289	89,384	140 bp	88,736	99.51%	27,517	100.00%
NOR10b	12,124,593	87,289	139 bp	86,776	99.14%	26,678	100.00%
NO12	8,283,891	62,500	133 bp	61,982	99.51%	18,289	100.00%
NO13	66,686,455	472,565	141 bp	469,257	99.39%	146,600	100.00%
NO14	7,998,411	55,964	143 bp	55,584	99.69%	17,721	100.00%
NO18	10,511,134	68,933	152 bp	68,474	98.83%	23,161	100.00%
NO19	7,899,227	54,381	145 bp	54,233	99.92%	18,049	70.02%
NO20	10,821,735	74,241	146 bp	73,705	99.71%	22,991	90.61%
NO21	10,693,294	74,115	144 bp	73,94	99.91%	24,153	70.02%
NO22	8,494,096	56,629	150 bp	56,225	99.66%	18,504	70.02%
NO24	8,743,766	57,528	152 bp	57,138	99.21%	19,401	100.00%
NO25	5,323,586	33,312	160 bp	32,129	99.77%	11,628	70.02%

Table 4.4.1.3 | Reads and coverage data for NGS sequencing.

On average, excluding data from sample NO13 that are significantly different from those concerning the other samples, 9,642,517.4 bases were read for sample producing a mean of

62,511 reads, 99.45% of which on target. Focusing on base coverage data, each base was read 21,207.8 and uniformity of base coverage was 89.51%. As stated, every sample was sequenced as a pool of amplicons of three fragments. The Torrent Server provided a diagram of the base coverage depth of the corresponding positions in the hypervariable region sequenced (Figure 4.4.1.4). In general, the first fragments had the highest values of coverage, whereas the second and the third showed lower values, likely due to the different nucleotides compositions of these regions, especially, the presence of the *c-stretch* in the second fragment.

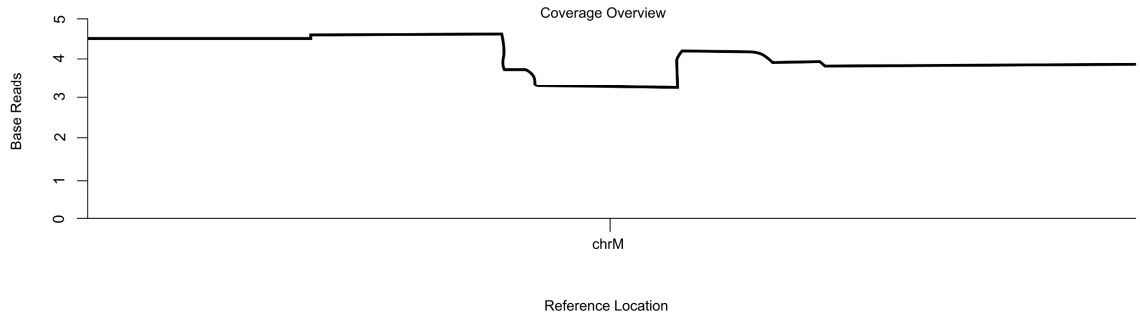


Figure 4.4.1.4 | Base coverage depth according to the base position on the HVR1 of the mitochondrial genome (example on sample NOR3a).

4.4.2. Sequences filtering

Figure 4.4.2.1 shows for each sample the reduction in the reads number at every step of the filtering process. The range of percentage reduction in the number of reads due to the quality and size filtering spans from 39.72% to 64.58%. It was observed that after the filtering the initial coverage depth difference among fragments incremented. On average, for the first fragment, 28,436.1 reads passed the filtering while just 4,971.1 were retained for the second and 5,843.2 for the third one. Moreover, focusing on the step where the reads from each sample were separated according to their sequencing primer, for the second and the third fragment an inequality can be noted between the numbers of the forward and reverse reads.

ID	NOR3a						NOR3b						NO19					
ION ID	Ion_1						Ion_2						Ion_3					
Loaded sequences	57,248						91,125						54,381					
Filtered sequences (nb)	30,365						53,281						24,338					
Filtered sequences (%)	53.04%						58.47%						44.75%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	14,086	11,379	1	1,108	22	3,593	176	24,909	18,541	5	3,041	452	5,970	363	9,300	7,575	1	6,743	15	484	220
Concatenated	25,465		1,109		3,615			43,450		3,046		6,422			16,875		6,744		499		
Clipped Primers	25,463		1,105		3,615			43,449		3,042		6,415			16,871		6,741		499		
Collapsed	1,562		223		444			1,241		424		590			1,417		1,126		81		
Counts ≥ 5	143		36		23			230		26		60			103		57		6		

ID	NO21						NO25						NOR8					
ION ID	Ion_4						Ion_5						Ion_6					
Loaded sequences	74,115						33,312						76,465					
Filtered sequences (nb)	37,814						21,514						30,373					
Filtered sequences (%)	51.02%						64.58%						39.72%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	15,779	13,293	2	7,483	7	837	413	9,211	8,38	0	3,073	279	309	262	11,344	9,486	28	4,321	49	4,780	365
Concatenated	29,072		7,485		844			17,591		3,073		588			20,830		4,349		4,829		
Clipped Primers	29,070		7,481		843			17,590		3,073		588			20,828		4,339		4,826		
Collapsed	863		728		107			435		404		96			2,020		1,076		975		
Counts ≥ 5	179		67		8			120		21		11			272		148		90		

ID	NOR9						NOR10a						NOR10b					
ION ID	Ion_7						Ion_8						Ion_9					
Loaded sequences	60,426						89,384						87,289					
Filtered sequences (nb)	28,921						38,169						38,645					
Filtered sequences (%)	47.86%						42.70%						44.27%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	11,037	9,788	6	2,161	6	5,598	325	14,872	12,616	5	3,099	570	6,604	403	16,392	12,607	31	3,619	53	5,529	414
Concatenated	28,825		2,167		5,604			27,488		3,104		7,174			28,999		3,650		5,582		
Clipped Primers	28,823		2,161		5,601			27,480		3,093		7,171			28,996		3,637		5,577		
Collapsed	998		357		348			2,088		545		845			2,649		603		577		
Counts ≥ 5	174		21		26			210		46		57			233		45		25		

ID	NO12						NO13						NO14					
ION ID	Ion_10						Ion_11						Ion_12					
Loaded sequences	62,500						218,809						55,964					
Filtered sequences (nb)	26,820						38,169						31,401					
Filtered sequences (%)	42.91%						46.30%						56.11%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	10,033	8,260	1	3,003	5	5,327	191	87,460	72,730	1,612	25,126	701	29,568	1612	12,351	9,578	0	2,765	3	6,436	268
Concatenated	18,293		3,004		5,332			160,190		26,738		30,269			21,929		2,765		6,439		
Clipped Primers	18,293		3,001		5,332			160,172		26,693		30,248			21,297		2,758		6,438		
Collapsed	1,092		507		769			5,071		2,127		2,550			409		284		434		
Counts ≥ 5	130		56		45			515		227		225			73		17		37		

ID	NO18						NO20						NO22					
ION ID	Ion_13						Ion_14						Ion_15					
Loaded sequences	68,933						74,241						56,629					
Filtered sequences (nb)	39,195						37,680						30,835					
Filtered sequences (%)	56.86%						50.75%						54.45%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	16,644	14,242	3	4,145	10	3,880	271	17,676	13,021	2	2,558	86	4,112	225	16,400	11,033	0	2,420	110	654	218
Concatenated	30,886		4,148		3,890			30,697		2,560		4,198			27,433		2,420		764		
Clipped Primers	30,881		4,143		3,890			30,694		2,550		4,196			27,429		2,420		762		
Collapsed	893		314		424			1,365		199		413			1425		194		123		
Counts ≥ 5	173		17		28			163		9		26			173		12		5		

ID	NO24					
ION ID	Ion_16					
Loaded sequences	57,528					
Filtered sequences (nb)	29,840					
Filtered sequences (%)	51.87%					
	1fw	1rv	2fw	2rv	3fw	3rv	NA
Splitted	10,316	8,344	173	3,148	2,443	5,050	366
Concatenated	18,660		3,321		7,493		
Clipped Primers	18,655		3,300		7,490		
Collapsed	1,189		433		516		
Counts ≥ 5	85		15		53		

Figure 4.4.2.1 | Sequences filtering.

4.4.3. Authentication of the mtDNA data

In this study, the following procedures and the rigorous criteria used to estimate the reliability of the aDNA results, have allowed to exclude any modern contamination and certify the authenticity of the ancient data with a high degree of confidence:

- *(i)* rigorous laboratory precautions for ancient DNA study to avoid contamination with modern DNA were followed (see Ancient DNA procedures);
- *(ii)* Samples were collected from freshly excavated archaeological site in virtually modern human 'DNA-free' conditions as described above (see Samples for genetic analysis), a circumstance that has been suggested to facilitate the discrimination between endogenous and contaminant DNA (Sampietro et al., 2006; Pruvost et al., 2007);
- *(iii)* no contamination was observed in all blank extractions and negative controls included in each reaction;
- *(iv)* all the ancient samples screened with the L15996-H16401 primer pair yielded no amplification products, indicating the absence of intact modern exogenous DNA;
- *(v)* all HVS-I sequences obtained from Novilara samples showed different haplotypes from those of operators involved in this study (Table 4.4.3.1);
- *(vi)* aDNA data was considered as genuine whenever a clear sequence was reproduced in all the overlapping portion of each adjacent fragment;
- *(vii)* no recurrent mutations were highlighted in the sequences obtained from ancient individuals, excluding a systematic exogenous contamination;
- *(viii)* DNA extraction and amplification were performed twice in a subset of Novilara samples (n = 4), starting from different bones from the same individuals. The HVS-I mtDNA sequences obtained by Sanger sequencing experiment confirmed the same haplotype derived from NGS sequencing reaction;
- *(ix)* the phylogenetic consistency of the haplotypes and matching haplogroup assignments of both HVR-I data and coding region SNPs, were indicative of the robustness of the mtDNA typing approach presented here;
- *(x)* the assay (PCR coupled with amplicon sequencing in NGS) has been shown to be highly sensitive for sequencing limited DNA amounts and to analyse biological mixtures of samples, allowing to detect low-level variants (Berglund et al., 2011). Moreover, this approach was useful to study damage patterns and point out contaminations from exogenous sources, by means of deep coverage data (Palencia-Madrid and de Pancorbo, 2015).

Researcher	HVSI range	HVSI Haplotype based on rCRS	Haplogroup based on HaploGrep	Overall Quality *
M1	15997-16409	16126C 16189C 16294T 16296T	T2+16189	1.000
M2	15997-16409	16183C 16189C 16194C 16195C	B4a1c3	0.808
M3	15997-16409	16126C 16153A 16183C 16189C 16294T 16296T	T2e1a1b	1.000
M4	15997-16409	16134T 16356C	U4a1	1.000

Table 4.4.3.1 | HVS-I motifs of the researchers who had been in contact with the ancient samples during the archaeological excavation and the laboratory work. * Overall Quality: 1 – 0.9 the haplogroup assignment is quite reliable 0.9 – 0.7 the haplogroup assignment is more accurate < 0.7 the haplogroup assignment is not very reliable.

4.4.4. Intra-population analysis

4.4.4.1. Haplogroup assignment

HVS-I mitochondrial consensus sequences were obtained in 10 out of the 27 ancient specimens (Table 4.4.4.1.1, S-Figure 4.4.4.1.2 in *Appendix I*), which represents an overall success rate of 37.03% for extraction and sequencing of mtDNA. The remaining 17 samples were excluded from subsequent analyses because yielded no amplification products (n = 11) or produced ambiguous sequence results (n = 6).

Sample	Manipulators[3]	rCRS position	HVS-I haplotype (NGS)	HVS-I haplotype (Sanger)	SNP in coding region	Haplogroup
NOR3a[1]	M1, M2, M3	16024-16383	rCRS	rCRS	7028C	H*
NOR3b[1]	M1, M2, M3	16024-16383	16291T	16291T	7028C	H*
NOR8[1]	M2, M4	16024-16383	16069T, 16126C, 16362C	16069T, 16126C, 16362C	3010A, 4216C	J1
NOR10a	M2, M4	16024-16261	16222T	-	7028C	H*
NOR10b[1]	M1, M2, M4	16024-16383	16192T, 16298C	16192T, 16298C	7028C	H*
NO12	M2, M4	16024-16383	16069T, 16126C	-	3010A, 4216C	J1
NO13	M2, M3	16024-16383	rCRS	-	3010A	H1*
NO19[2]	M2, M4	16024-16383	16224C, 16311C	-	-	K
NO20[2]	M2, M3	16024-16383	16356C	-	-	HV1
NO21[2]	M2, M4	16024-16383	16145A, 16234T, 16270T	-	-	H1

Table 4.4.4.1.1 | mtDNA sequences of samples from Novilara site. mtDNA haplotypes were numbered according to the rCRS (Andrews et al., 1999). [1]Samples analysed two times and sequenced with both Sanger and NGS methods, [2]Haplogroup predicted with Haplogrep2 software, [3]for HVS-I motifs of the researchers see Table 4.4.3.1.

A partial HVS-I consensus sequence was obtained for NOR10a (np 16024-16261), making this sample useless for the population genetic analysis. Haplogroups preliminarily inferred by HVS-I mutation motif were confirmed by the genotyping of 22 mtDNA SNPs (see Table 4.4.4.1.1). However, in three samples the multiplex amplification failed (NO19, NO20 and NO21). By combining sequence and genotyping analyses, the investigated samples were

classified as belonging to six different mtDNA lineages: H* (NOR3a, NOR3b, NOR10a and NOR10b), J1 (NOR8 and NO12), K (NO19), H1* (NO13), H1 (NO21) and HV1 (NO20) (see Table 4.4.4.1.1). All these lineages have been reported to be typical of West Eurasian region (Richards et al., 2000). It is worth noting that the four samples belonging to the same paragroup H* were found in shared graves (grave 155 for NOR10a and NOR10b and grave 171 for NOR3a and NOR3b), but carrying different mutation motifs.

4.4.4.2. Summary and population differentiation statistics

The genetic and standard diversity indexes estimated for the Piceni population are shown in Table 4.4.4.2.1a, and Table 4.4.4.2.1b, together with the same values calculated for other present-day Italian populations.

Population	*Area*	*n*	*Haplot. Diff. (K)*	*Gene diversity (h) ± sd*	*MNPD (π) ± sd*	*Nucleot.Div. (πN) ± sd*
Cuneo	1	40	29	0.9654 +/- 0.0193	3.9979 +/- 2.0412	0.0111 +/- 0.0062
Como	1	39	36	0.9946 +/- 0.0078	5.0073 +/- 2.4867	0.0139 +/- 0.0076
Brescia	1	40	30	0.9731 +/- 0.0158	5.3081 +/- 2.6174	0.0147 +/- 0.0080
Savona/Genova	1	43	28	0.9336 +/- 0.0312	4.9583 +/- 2.4602	0.0137 +/- 0.0075
Aviano	2	29	28	0.9886 +/- 0.0109	5.8323 +/- 2.8603	0.0162 +/- 0.0088
Vicenza	2	40	32	0.9808 +/- 0.0129	5.3911 +/- 2.6538	0.0149 +/- 0.0081
Treviso	2	39	29	0.9690 +/- 0.0173	5.5859 +/- 2.7408	0.0155 +/- 0.0084
Udine	2	51	43	0.9890 +/- 0.0081	5.4284 +/- 2.6580	0.0158 +/- 0.0086
Bologna	3	100	65	0.9671 +/- 0.0119	4.7369 +/- 2.3368	0.0138 +/- 0.0075
Modena	3	43	32	0.9557 +/- 0.0239	4.0201 +/- 2.0479	0.0111 +/- 0.0063
Grosseto/Siena	4	36	21	0.8873 +/- 0.0478	4.5225 +/- 2.2776	0.0125 +/- 0.0070
Pistoia	4	16	15	0.9917 +/- 0.0254	4.7422 +/- 2.4478	0.0131 +/- 0.0076
Casentino	4	120	76	0.9782 +/- 0.0072	4.8623 +/- 2.3879	0.0135 +/- 0.0073
Murlo	4	86	61	0.9759 +/- 0.0098	4.8668 +/- 2.3965	0.0135 +/- 0.0073
Volterra	4	113	57	0.9539 +/- 0.0134	4.3733 +/- 2.1768	0.0121 +/- 0.0066
Terni	4	31	40	0.9753 +/- 0.0137	4.5520 +/- 2.2734	0.0126 +/- 0.0070
Macerata	5	39	35	0.9923 +/- 0.0079	5.2277 +/- 2.5821	0.0145 +/- 0.0079
Foligno	5	40	27	0.9785 +/- 0.0197	5.5157 +/- 2.7255	0.0153 +/- 0.0084
Ascoli Piceno	5	53	31	0.9838 +/- 0.0109	4.9507 +/- 2.4619	0.0137 +/- 0.0075
Piceni from Novilara	5	9	8	0.9722 +/- 0.0640	3.1549 +/- 1.8003	0.0087 +/- 0.0056
Ancona	5	70	52	0.9602 +/- 0.0180	4.9094 +/- 2.4204	0.0136 +/- 0.0074
L'Aquila	6	25	23	0.9933 +/- 0.0134	4.6356 +/- 2.3522	0.0128 +/- 0.0072
Benevento	6	36	33	0.9905 +/- 0.0114	5.5589 +/- 2.7339	0.0154 +/- 0.0084
Cosenza/Catanzaro/ Crotone	6	37	30	0.9670 +/- 0.0224	5.4337 +/- 2.6772	0.0150 +/- 0.0082
Belvedere	6	50	36	0.9829 +/- 0.0086	5.0392 +/- 2.4888	0.0139 +/- 0.0076
Campobasso	6	37	29	0.9775 +/- 0.0148	4.4227 +/- 2.2322	0.0122 +/- 0.0068
Matera	6	36	34	0.9952 +/- 0.0087	5.9040 +/- 2.8856	0.0164 +/- 0.0089
Lecce	6	39	35	0.9946 +/- 0.0071	6.2363 +/- 3.0260	0.0173 +/- 0.0093
Sanniti	6	50	30	0.9682 +/- 0.0117	5.6139 +/- 2.7399	0.0155 +/- 0.0084
Lucera	6	60	42	0.9757 +/- 0.0110	4.9963 +/- 2.4631	0.0146 +/- 0.0079
Agrigento	7	42	28	0.8931 +/- 0.0457	4.2008 +/- 2.1285	0.0116 +/- 0.0065
Catania	7	37	30	0.9670 +/- 0.0224	5.7843 +/- 2.8312	0.0160 +/- 0.0087
Ragusa/Siracusa	7	39	28	0.9690 +/- 0.0167	5.2110 +/- 2.5762	0.0144 +/- 0.0079
Trapani	7	40	32	0.9808 +/- 0.0129	4.9537 +/- 2.4618	0.0145 +/- 0.0080
Sardinia	8	233	118	0.9466 +/- 0.0117	4.3773 +/- 2.1704	0.0127 +/- 0.0070

Table 4.4.4.2.1a | Standard and genetic diversity indexes estimated for the Piceni population and for the present-day Italian populations.

Italian macro-area		n	Haplot. Diff. (K)	Gene diversity (h) ± sd	MNPD (π) ± sd	Nucleot.Div. (πN) ± sd
IT_Area 1	-	162	101	0.9689 +/- 0.0093	4.8581 +/- 2.3819	0.0134 +/- 0.0073
IT_Area 2	-	163	107	0.9824 +/- 0.0052	5.5696 +/- 2.6891	0.0162 +/- 0.0087
IT_Area 3	-	143	88	0.9625 +/- 0.0111	4.5315 +/- 2.2421	0.0132 +/- 0.0072
IT_Area 4	-	371	177	0.9673 +/- 0.0062	4.7043 +/- 2.3087	0.0130 +/- 0.0070
IT_Area 5	-	233	150	0.9770 +/- 0.0062	4.9766 +/- 2.4293	0.0138 +/- 0.0074
IT_Area 6	-	370	218	0.9838 +/- 0.0035	5.3461 +/- 2.5853	0.0156 +/- 0.0083
IT_Area 7	-	158	106	0.9578 +/- 0.0125	5.0507 +/- 2.4654	0.0147 +/- 0.0079
IT_Area 8	-	233	118	0.9466 +/- 0.0117	4.3773 +/- 2.1704	0.0127 +/- 0.0070

Table 4.4.4.2.1b | Standard and genetic diversity indexes estimated for the Piceni population and for the present-day Italian populations.

Interestingly, the nucleotide diversity of ancient Piceni (0.0087 ± 0.0056) resulted to be lower (albeit not significantly) than the diversity indices reported for modern Italian populations. Given the small sample size of the Piceni, which could mislead the interpretation of the intra-population statistics, a resampling procedure was performed. We randomly extracted, without replacement, 1000 subsamples of 9 individuals each, from all the 34 Italian extant populations of our dataset and then we re-calculated the nucleotide diversity for each subset. The resulting distributions were then compared with the value observed for the ancient Novilara sample. Although with no statistical robustness, by the resampling procedure we confirmed the results. In fact, the nucleotide diversity values of the ancient Piceni always falls within the 1st quartile or even outside (Lecce) the distributions obtained through the resampling of the extant Italian populations (Figure 4.4.4.2.2).

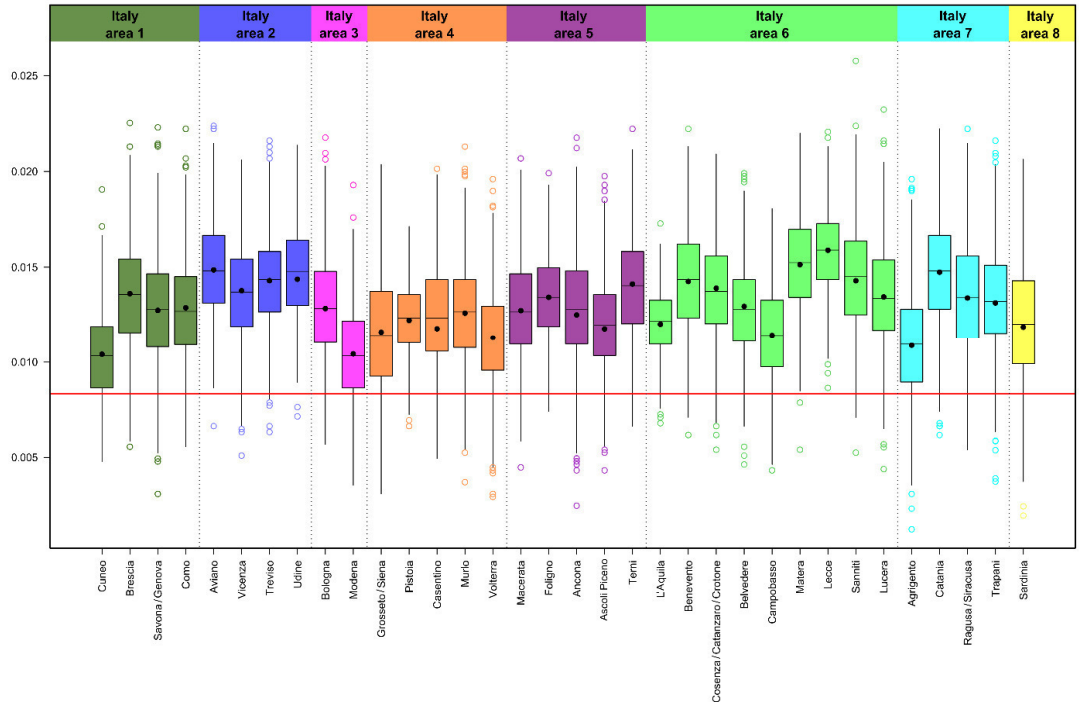


Figure 4.4.4.2.2 | Distribution of the nucleotide diversity (π) between 1000 subsamples extracted from each of 34 Italian populations. The red line indicates the observed nucleotide diversity value of the Novilara sample.

4.4.4.3. Kinship assessment

The autosomal analysis returned partials profiles for a number of the tested samples: NOR10a and NOR10b (grave 155) being the most successful ones, followed by NOR3a and NOR3b (grave 171).

The samples of the grave 155 allowed the amplification of an almost full profile with InDels kit (Table 4.4.4.3.1) and a fairly complete Globalfiler kit STR profile (Table 4.4.4.3.2), resulting in a 93.3/60 percent of completeness (DIPplex/Globalfiler) for NOR10A and 83.3/44 percent of completeness for NOR10B. Both profiles showed a different allele combination in comparison with the other successfully amplificated samples and with respect to the laboratory staff´s profiles. The comparison of the Globalfiler results from the two independent amplification experiments confirmed the obtained profiles and pointed out three drop-outs in two loci for NOR10a (TH01 and FGA) and one locus for NOR10b (D22S1045) (Table 4.4.4.3.2). For the subjects of the grave 171 no results were obtained with the Globalfiler analysis, while we achieved a partial profile through DIPplex kit. The different completeness percentage observed among InDel and STR results was to be expected due to the shorter amplicon length sported by InDels when compared to what STR typing may offer, including the miniSTR mode amplification. It has to be noted that, in parallel to the severe degradation of the samples, the issue of low copy number DNA may also contribute to the observed results. This lead to a certain degree of randomness in the observed results compared to the expected, serving as an explanation for the partial amplification obtained in one assay, but not in the other, and for the absence of successful replication for these profiles.

The amelogenin locus indicated that NOR10a, NOR10ba and NOR3a subjects were females. Due to the failure of the Globalfiler reaction, the only DIPplex analysis for the NOR3b sample does not allow us to confirm with a high degree of confidence the sex of this individual. In fact, the presence of a peak (sz 79.85; ht 150) at the Y position, completely unbalanced in comparison with the X one (sz 76.96; ht 1612), should be confirmed by additional analyses. Accordingly, we do not feel confident to assign with certainty the sex of the individual NOR3b, for which, the archaeological and anthropological data suggested that is a male. Anyway, in the other three samples analysed with autosomal markers, the sex determination through archaeological, osteological, and genetic analysis showed a complete concordance.

DIPplex kit marker	rs number	Grave 155		Grave 171	
		NOR10A	NOR10B	NOR3A	NOR3B
Amel	N/A	XX	XX	XX	?
D77	rS1611048	11	11	12	11
D45	rS2307959	12	12	12	22
D131	rS1611001	22	12	-	22
D70	rS2307652	12	12	11	22
D6	rS1610905	11	11	-	-
D111	rS1305047	22	11	-	-
D58	rS1610937	11	22	-	-
D56	rS2308292	22	12	-	-
D118	rS16438	11	-	11	12
D92	rS201771066	12	11	22	12
D93	rS150042219	22	12	12	11
D99	rS2308163	12	22	-	1n
D88	rS8190570	n2	12	-	-
D101	rS2307433	11	11	-	-
D67	rS1305056	12	11	-	-
D83	rS2308072	12	-	22	12
D114	rS2307581	11	22	12	11
D48	rS28369942	11	12	12	11
D124	rS6481	22	12	11	12
D122	rS8178524	11	11	-	-
D125	rS16388	12	12	-	-
D64	rS397832668	22	22	-	22
D81	rS17879936	22	-	-	-
D136	rS16363	12	22	12	11
D133	rS2067235	22	12	11	22
D97	rS17238892	11	22	11	11
D40	rS146044344	-	-	-	-
D128	rS2307924	22	11	-	11
D39	rS17878444	11	11	-	22
D84	rS3081400	-	-	-	-

Table 4.4.4.3.1 | Autosomal insertion-deletion (InDel) profile of Novilara samples (NOR10a; NOR 10b; NOR3a; NOR3b) buried in two bisome-graves. 1=deletion; 2=insertion.

Globalfiler kit marker	Grave 155	
	NOR10A	NOR10B
D3S1358	14, 18	16/ **16**
vWA	17, 18	18, **19**
D16S539	-	
CSF1PO	-	12
TPOX	-	-
Yindel	-	-
AMEL	X, X	X, X
D8S1179	10, 14	13, 16
D21S11	-	-
D18S51	-	-
DYS391	-	-
D2S441	11, 14	11.3/ **11.3**
D19S433	12, 13.2	14/ **14**
TH01	6/ **5.3, 6**	-
FGA	20/ **20, 25**	-
D22S1045	11, 17	11/ **11, 14**
D5S818	11	12
D13S317	10	-
D7S820	-	-
SE33	-	-
D10S1248	13, 14	13, 14/ **13, 14**
D1S1656	11	-
D12S391	17.3, 19	-
D2S1338	-	-

Table 4.4.4.3.2 | Autosomal STRs profile of Novilara samples (NOR10a; NOR 10b; NOR3a; NOR3b) buried in two bisome-graves. No result was obtained for NOR3a and NOR3b samples. Obtained results of independent PCR reactions are reported in bold.

In order to give a clear statistical support to the Novilara individuals' kinship investigations, the genealogical relationships between the two pairs of individuals buried in the same grave (i.e. NOR3a-NOR3b for grave 171 and NOR10a-NOR10b for grave 155) were estimated using the software Familias, Version 3.2.1 (Egeland et al., 2000; Kling et al., 2014). Considering the mtDNA data, according to which the 4 samples belong to the same haplogroup, but with different haplotypes, we inferred the lack of maternal relationship between the two pairs of subjects. Blind Search module, a new tool in this version of Familias, was run to perform an unspecific relationship search for a set of persons with some DNA data. The Direct-Match, Parent-Child, Siblings, Half Siblings, Cousins, 2nd Cousins relationships were tested. For NOR10a and NOR10b, the LRs (likelihood ratio) were calculated using both available autosomal STRs (8 loci) and InDels (21) markers, while for NOR3a and NOR3b only the 14 obtained InDels were considered (for the STRs allele frequencies: Presciuttini et al., 2006; Previderè et al., 2013; InDels frequencies: QIAGEN© 2010 Population Data for analysis of results from the Investigator DIPplex Kit). For both pairs of individuals, LR values were below zero (Table 4.4.4.3.3a, Table 4.4.4.3.2b), so these results are not supporting the different tested kinship hypotheses. Moreover, the values are gradually increasing inversely to the degree of kinship, as expected.

STRs (8 loci) + InDels (21 markers)			
Individual 1	**Individual 2**	**Kind of relationship**	**LR**
NOR10a	NOR10b	2nd Cousins	0.757948
NOR10a	NOR10b	Cousins	0.214652
NOR10a	NOR10b	Half-siblings	0.0106976
NOR10a	NOR10b	Siblings	9.58E-01
NOR10a	NOR10b	Parent-Child	0
NOR10a	NOR10b	Direct-match	0

Table 4.4.4.3.3a | Relationships established between individuals NOR10a and NOR10b whit software Familias, Version 3.2.1.

InDels (14 markers)			
Individual 1	**Individual 2**	**Kind of relationship**	**LR**
NOR3a	NOR3b	2nd Cousins	0.981487
NOR3a	NOR3b	Cousins	0.845878
NOR3a	NOR3b	Half-siblings	0.526414
NOR3a	NOR3b	Siblings	0.0170355
NOR3a	NOR3b	Parent-Child	0
NOR3a	NOR3b	Direct-match	0

Table 4.4.4.3.3b | Relationships established between individuals NOR3a-NOR3b whit software Familias, Version 3.2.1.

The genetic data presented here provide the first attempt to support archaeological hypotheses based on material culture data (such as graves distribution and manufacture styles) in the interpretation of funerary practices and kinship relationships. These data derive from a limited number of genetically analysed samples, with regard to those originally present in the sampling design of the project (see S-Figure 4.4.4.1.2 in *Appendix I*), selected on the basis of archaeological data in particular areas of the necropolis, as potentially linked to familiar or clan-related groups (see Delpino et al., 2016). Unfortunately, the 10 individuals for which the mitochondrial data passed the stringent criteria adopted in this study were located in burials physically scattered across the necropolis (see S-Figure 4.4.4.1.2 in *Appendix I*). This random distribution of the genetic results did not allow speculating about specific funerary practices that occurred in contiguous groups of graves. However, it can be only state that the mitochondrial data seem to exclude a matrilineal relationship between them (see S-Figure 4.4.4.1.2 in *Appendix I*).

Moreover, autosomal analyses were provided only for 4 individuals buried in the two bisome graves (S-Figure 4.4.4.1.2 in *Appendix I*). Regarding the burial 171, the two adult individuals identified as a woman and probably a man, were buried in the same burial place in two different times: first was buried the "individual b" (sample NOR3b – the alleged man) and then, right above the first one, was buried the "individual a" (sample NOR3a – a women). Although buried together, autosomal data did not support a family relationship between them, a situation that could lead to hypothesize the reuse of the same burial place for two subjects unrelated or related-in law. The particular funerary set discovered in this grave pushes archaeologists to believe that individuals NOR3a and NOR3b were somehow connected together. The grave goods of this burial belong almost completely to the "individual a", a woman, and they are constituted, among the vases, by typical female brooches, and other female adornments. However, upon the women body, there were also some fragments of typical male brooches, probably wrapped in a kind of a textile such as a shroud. It seems that when the first grave was re-opened to place inside also the female corpse, the first funerary set was collected and put up on the second individual, along with her proper personal grave goods. The autosomal data obtained from the individuals of the bisome grave 155 (containing two females buried simultaneously), did not support the idea of a kinship between them. The burials of two adult females that, apparently, did not show signs of violent death or infectious diseases (at least for those diseases that leave signs on the bones) leaves open new scenarios and interpretation.

4.4.5 Inter-population diversity

A Multidimensional scaling (MDS) analysis comparing the HVS-I mitochondrial variability (np 16024-16383) of the Novilara individuals, 12 European prehistoric populations and 1833 present-day Italians was carried out to provide a two-dimensional plot of the Fst genetic distances matrix. The results of genetic distance indexes calculated from three different datasets, are shown in Table 4.4.5.1a; b; c.

	Piceni_a	*IT_area1*	*IT_area2*	*IT_area3*	*IT_area4*	*IT_area5*	*IT_area6*	*IT_area7*	*IT_area8*
Piceni_a	*	0	0	0	0	0	0	0	0
IT_area1	-0.03057	*	0.00357	0.00217	0.00139	0.00033	0.00247	0.00269	0.00397
IT_area2	-0.02137	0.00356	*	0.00598	0.00285	0.00109	0.00023	0	0.00529
IT_area3	-0.02293	0.00216	0.00594	*	0.00219	0.00021	0.00507	0.00583	0.00639
IT_area4	-0.02032	0.00138	0.00284	0.00218	*	0.00181	0.002	0.00238	0.00346
IT_area5	-0.02114	0.00033	0.00109	0.00021	0.00181	*	0.00083	0.00214	0.00247
IT_area6	-0.01963	0.00247	0.00023	0.00504	0.002	0.00083	*	0.00019	0.00372
IT_area7	-0.02134	0.00269	-0.00126	0.0058	0.00237	0.00214	0.00019	*	0.00429
IT_area8	-0.01987	0.00395	0.00526	0.00635	0.00345	0.00247	0.0037	0.00427	*

Table 4.4.5.1a | Fst Matrix used for MDS in Figure 4.4.5.1.1a. Above diagonal: Fst values (Slatkin); below diagonal: Fst values (not normalised). For the list of populations see Table 4.3.2.1.2.

	PIa	*ANC*	*MAC*	*FOL*	*ASPI*	*VOL*	*MUR*	*CAS*	*GRSI*	*PIS*	*TER*
PIa	*	0	0	0	0	0	0	0	0	0	0
ANC	-0.02192	*	0	0	0	0	0	0.00174	0.00439	0	0.00044
MAC	-0.01589	-0.00627	*	0.00504	0	0	0.00015	0	0.00312	0	0
FOL	-0.02219	-0.00321	0.00501	*	0	0.00905	0.01027	0.00466	0.01574	0	0.00172
ASPI	-0.02172	-0.00619	-0.00382	-0.00363	*	0.00059	0	0	0.00823	0	0
VOL	-0.01602	-0.00063	-0.00155	0.00897	0.00059	*	0.00278	0.00316	0.00873	0	0.00737
MUR	-0.02476	-0.00031	0.00015	0.01017	-0.00023	0.00277	*	0.00339	0	0	0
CAS	-0.01556	0.00173	-0.00255	0.00464	-0.00013	0.00315	0.00338	*	0.002	0	0.00212
GRSI	-0.02481	0.00437	0.00311	0.01549	0.00816	0.00866	-0.0035	0.00199	*	0	0.00254
PIS	-0.01433	-0.02033	-0.02361	-0.01014	-0.01441	-0.0154	-0.01225	-0.01416	-0.00701	*	0
TER	-0.01809	0.00044	-0.00041	0.00172	-0.00347	0.00731	-0.00135	0.00212	0.00254	-0.00981	*

Table 4.4.5.1b | Fst Matrix used for MDS in Figure 4.4.5.1.1b. Above diagonal: Fst values (Slatkin); below diagonal: Fst values (not normalised). Population code: PIa, Piceni from Novilara; ANC, Ancona; MAC, Macerata; FOL, Foligno; ASPI, Ascoli Piceno; VOL, Volterra; MUR, Murlo; CAS, Casentino; GRSI, Grosseto Siena; PIS, Pistoria; TER, Terni.

	PIC	*NUR*	*ETR*	*SPA*	*CRE*	*TK*	*SPA*	*GER*	*CAT*	*YAM*	*PRZ*	*SCYR*	*SCYU*
PIC	*	0.00036	0.01033	0	0	0.06164	0	0.08728	0	0.02478	0.00005	0.02308	0.011
NUR	0.00036	*	0.06656	0.05347	0.01274	0.01566	0.02986	0.18364	0.02558	0.05378	0.02629	0.07861	0.08785
ETR	0.01022	0.06241	*	0.02931	0.06565	0.0699	0.03231	0.13916	0.04552	0.03119	0.05801	0.05167	0.07544
SPA	-0.06887	0.05076	0.02847	*	0.02765	0.27682	0	0.10619	0	0.02771	0.05115	0.03752	0.01128
CRE	-0.00108	0.01258	0.0616	0.0269	*	0.02819	0.01619	0.04465	0.02418	0.03	0.02733	0.03977	0.04464
TK	0.05806	0.01542	0.06533	0.2168	0.02742	*	0.07241	0.33028	0.0333	0.10222	0.01049	0.0749	0.06756
SPA	-0.02481	0.029	0.0313	-0.02826	0.01593	0.06752	*	0.08042	0.00611	0.00127	0.0273	0.00394	0.00986
GER	0.08028	0.15515	0.12216	0.096	0.04274	0.24828	0.07443	*	0.05398	0.12866	0.14201	0.07647	0.08089
CAT	-0.03658	0.02494	0.04354	-0.02464	0.02361	0.03223	0.00608	0.05122	*	0.03331	0.01494	0.02953	0.00578
YAM	0.02418	0.05104	0.03025	0.02697	0.02912	0.09274	0.00127	0.11399	0.03223	*	0.04824	0	0.04796
PRZ	0.00005	0.02561	0.05483	0.04866	0.0266	0.01038	0.02657	0.12435	0.01472	0.04602	*	0.02538	0.00757
SCYR	0.02256	0.07288	0.04913	0.03616	0.03825	0.06968	0.00393	0.07104	0.02868	-0.00283	0.02475	*	0.00453
SCYU	0.01088	0.08075	0.07015	0.01116	0.04273	0.06329	0.00976	0.07484	0.00575	0.04577	0.00751	0.00451	*

Table 4.4.5.1c | Fst Matrix used for MDS in Figure 4.4.4.6.1. Above diagonal: Fst values (Slatkin); below diagonal: Fst values (not normalised). For population code see Table 4.3.2.1.1.

4.4.5.1 Genetic comparison whit modern populations

A first exploratory MDS was performed between Novilara samples and current Italian populations grouped in 8 macro-areas (see Table 4.4.4.2.1). The obtained plot (Figure 4.4.5.1.1a) showed the separation between Sardinia (area 8) and all the other populations of continental Italy and Sicily (areas 1, 2, 3, 4, 5, 6 and 7). In this context, Piceni from Novilara appears at the centre of the plot, occupying a position very close to the nearby Italian populations from the area 5 (Macerata, Foligno, Ancona, Ascoli Piceno and Terni).

As a further insight, a second MDS analysis was carried out with only the 10 populations belonging to the macro areas that in the previous plot were plaeced near to Piceni (area 4 and area 5). The new MDS plot (Figure 4.4.5.1.1b) suggest a certain degree of genetic continuity, between Piceni from Novilara and the present-day inhabitants of the Marche region: the nearest ones being Ancona and Macerata populations, followed by the Ascoli Piceno population, encompassing specimens collected in small towns of the Ascoli Piceno province (the "Piceni" of Montefortino, Castorano and Offida available from Brisighelli et al. (2012)).

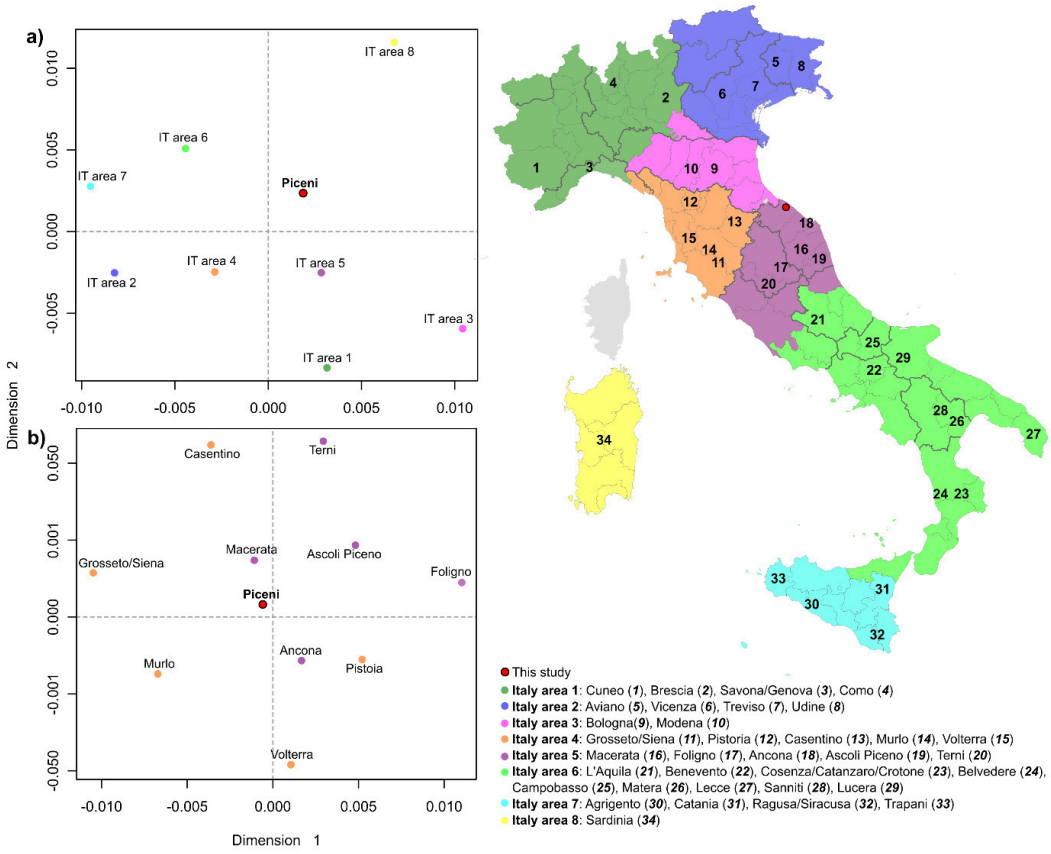


Figure 4.4.5.1.1 | Two-dimensional MDS plots of pairwise Fst values from HVS-I showing relationships among the 34 populations from continental Italy, Sicily, and Sardinia: a) MDS performed between individuals from Novilara and current Italian populations grouped in 8 macro-areas; b) MDS performed between ancient Piceni and 10 populations from central Italy areas (area 4 and area 5). The red circle represents the population from Novilara.

A second resampling procedure was performed to formally test for the higher genetic affinity of ancient Novilara group with the present-day inhabitants of Central Italy. To do so, 1000 sub-samples of N = 9, without replacement, were extracted from three Central Italian populations (Ascoli Piceno, Foligno and Macerata), three Northern Italian populations (Brescia, Udine and Savona/Genova) and three Southern Italian populations (Lecce, Matera, Ragusa/Siracusa). Then, the Fst value was calculated for each of these subpopulations and the Ancona sample, which is the geographically closest population to Novilara included in the comparison dataset. The resulting distributions were then compared to the Fst value obtained between Novilara and Ancona. This approach allowed us to test for a model of genetic affinity represented by the comparison between Ancona and Central Italian populations, and of population differentiation represented by Ancona vs Northern and Southern Italian populations. The resampling procedure was performed using the script by Anagnostou et al. (2017).

The resampling procedure highlighted that the genetic distance between Novilara and Ancona falls adequately within the Fst distributions obtained with the genetic affinity rather than with the population differentiation model (Figure 4.4.5.1.2).

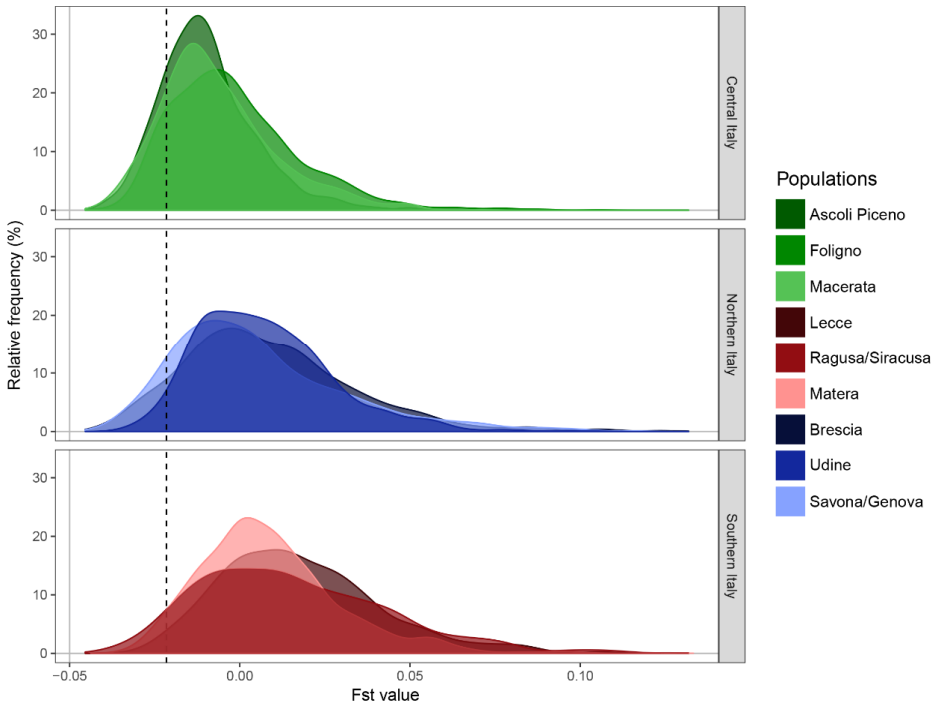


Figure 4.4.5.1.2 | Distribution of Fst genetic distances between 1000 subsamples extracted from each of nine Italian populations (three for each Northern, Central and Southern Italy) and Ancona. The dashed line indicates the observed Fst genetic distance value between the ancient Novilara sample and Ancona.

Overall, the obtained data indicate that the mtDNA diversity of Piceni of Novilara falls in the geographical cline of the mtDNA Italian genetic variability, which in general highlighted to retain a weaker genetic structure in the modern Italian population compared to Y-chromosome genetic diversity (Brisighelli et al., 2012; Boattini et al., 2013). Albeit preliminary and being aware that the mtDNA perspective may disclose only a part of the population history, our results seem to suggest that probably there was not such a strong reshuffling in the maternal genetic pool of the investigated area during historical periods, when, for instance, migrations due to Celts, Romans and Goths are attested. Indeed, as suggested by previous studies on present-day Italian population, the actual sex-biased genetic structure in Italy is possibly the result of different demographic histories for males and females, with the more homogenous pattern of mtDNA variability probably tracing back to more ancient times, and the Y-chromosome structure being instead shaped by more recent migration events (Boattini et al., 2013). If this genetic continuity involved other genomic loci remain to be elucidated.

4.4.5.2 Genetic comparison with ancient populations

In order to include the Piceni of Novilara within the genetic makeup of the European continent in the same temporal frame, we performed a third MDS plot based on the Fst genetic distances between Piceni and several ancient populations of the Bronze and Iron Age (Figure 4.4.5.2.1).

No particular geographic cluster has been observed from this comparison, but, interestingly, the Piceni resulted to be in the centre of the MDS plot indicating a genetic affinity with different ancient populations of continental Europe. This result could be probably due to the scarce availability of molecular data of ancient populations from this time frame, in particular from the Italian Peninsula. Moreover, given that some data originate from pioneering studies obtained with classical methods (cloning and Sanger sequencing), it will be desirable to acquire more genetic information through high resolution methodologies, as deep sequencing with NGS, in order to detect with more confidence exogenous contamination by modern DNA (Rizzi et al., 2012). These results are decisive incentives to make further research about Iron Age populations of Italy.

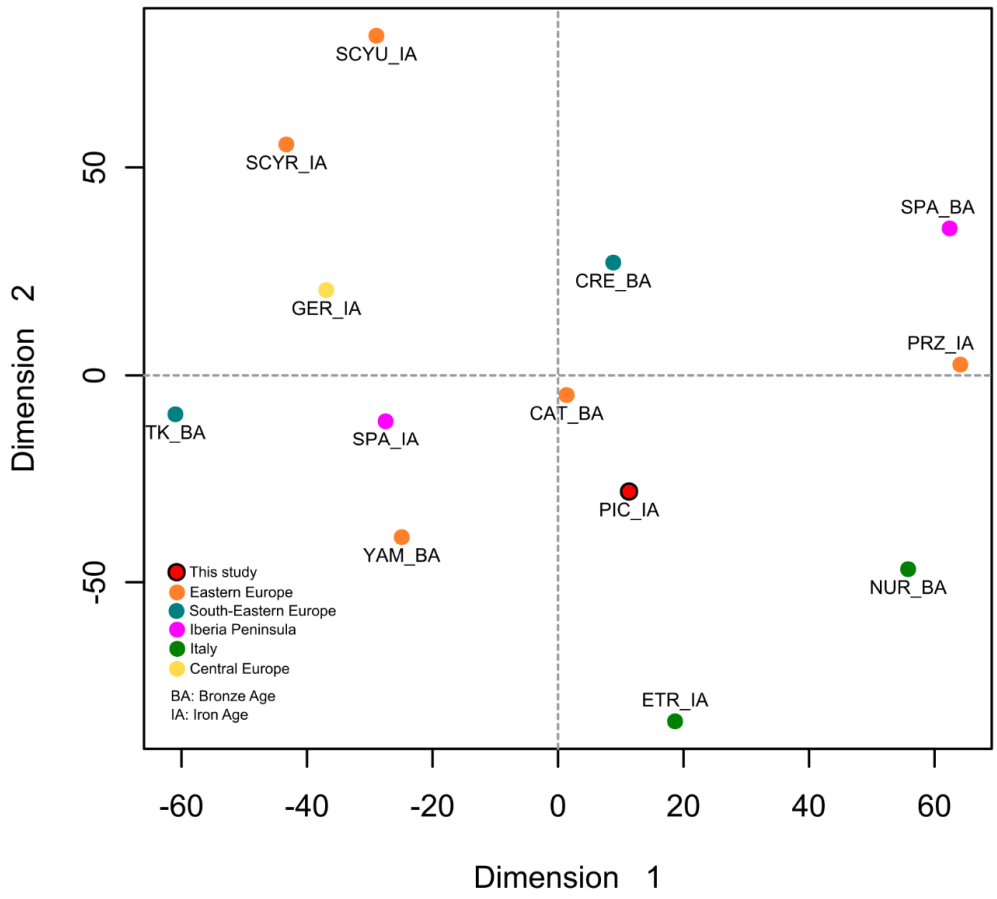


Figure 4.4.5.2.1 | Two-dimensional MDS plot of pairwise Fst values from HVS-I showing relationships among Novilara sample and 12 ancient populations from continental Europe (Bronze Age/Iron Age period). The red circle represents the Piceni from Novilara necropolis. Population codes are reported in Table 4.3.2.1.1.

4.5. Conclusion and Future Objectives

This study provides a preliminary characterization of the mitochondrial DNA variation of the Piceni from the Iron Age necropolis of Novilara (8th and 7th centuries BC) and contributes to enlarge the knowledge of the Italian populations of this period, for which few genetic data are currently available.

Despite being aware that more informative and reliable technologies are becoming progressively available and suitable for aDNA analysis (i.e. complete mitochondrial DNA capture), we retain that the methodologies and the strict criteria adopted in this study could constitute a good compromise between deep sequencing of entire mitochondrial or nuclear genomes, and a classical low-resolution approach constituted by cloning and Sanger sequencing of the HVS-I mtDNA, even now utilized (Krzewińska et al., 2015; Rivollat et al., 2015; Vai et al., 2015; Csősz et al., 2016; Kefi et al., 2016; Le Roy et al., 2016; Simòn et al.,

2016; Nikitin et al., 2017; Beau et al., 2017; Palencia-Madrid et al., 2017). The Ion semiconductor sequencing allows a better view of the molecular damage of the DNA fragments analysed. This method provides a high number of sequences, and thus, it is possible to observe a higher number of molecular modifications on them, obtaining a more precise picture of the damage of the molecule of DNA analysed. The apparent matrilineal genetic continuity between the ancient and modern populations in the region here analysed suggests that probably the different migratory events that involved this area did not influence the maternal gene pool of their inhabitants. If this continuity is maintained also for other loci remain to be elucidated.

This study also provides the first genetic data about the correlations between kinship and specific funerary traits in Novilara necropolis, as already analysed in previous studies (Le Roy et al., 2016). The few burials examined in this study, represented by two bisome burials, seem to not reveal a direct genetic relationship between the inhumates.

However, a greater richness of genetic data would be desirable in order to completely certify such conclusion. It will be advisable in the future to analyse more ancient samples from various necropolises of Piceni, belonging to different chronological period and localities, in order to better understand to what extent they can be considered a single population or smaller communities, which only partly recognized themselves in a wider organization (Carfagna, 2016).

References:

Achilli A, Olivieri A, Pala M, Metspalu E, Fornarino S, Battaglia V, Accetturo M, Kutuev I, Khusnutdinova E, Pennarun E, et al. 2007. Mitochondrial DNA Variation of Modern Tuscans Supports the Near Eastern Origin of Etruscans. Am. J. Hum. Genet. 80:759–768.

Anagnostou P, Capocasa M, Dominici V, Montinaro F, Coia V, Destro-Bisol G. 2017. Evaluating mtDNA patterns of genetic isolation using a re-sampling procedure: A case study on Italian populations. Ann. Hum. Biol. 44:140–148.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–65.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mito-chondrial DNA. Nat. Genet. 23:147.

Antonelli L, 2003. I Piceni. Corpus delle Fonti. La Documentazione Letteraria. Roma:L'Erma di Bretschneider.

Babalini C, Martínez-Labarga C, Tolk H-V, Kivisild T, Giampaolo R, Tarsi T, Contini I, Barać L, Janićijević B, Martinović Klarić I, et al. 2005. The population history of the Croatian linguistic minority of Molise (southern Italy): a maternal view. Eur. J. Hum. Genet. 13:902–912.

Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr W, Olaisen B. 1997. DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. Int. J. Legal Med. 110:175–6.

Barbujani G, Bertorelle G, Capitani G, Scozzari R. 1995. Geographical structuring in the mtDNA of Italians. Proc. Natl. Acad. Sci. U. S. A. 92:9171–5.

Beau A, Rivollat M, Réveillas H, Pemonge M-H, Mendisco F, Thomas Y, Lefranc P, Deguilloux M-F. 2017. Multi-scale ancient DNA analyses confirm the western origin of Michelsberg farmers and document probable practices of human sacrifice. Pinhasi R, ed-itor. PLoS One 12:e0179742.

Beinhauer KW. 1985. Untersuchungen zu den Bestattungsplätzen von Novilara (Provinz Pesaro und Urbino, Italien), Frankfurt am Main.

Berglund EC, Kiialainen A, Syvänen A-C. 2011. Next-generation sequencing technolo-gies and applications for human genetic history and forensics. Investig. Genet. 2:23.

Bertoncini S, Bulayeva K, Ferri G, Pagani L, Caciagli L, Taglioli L, Semyonov I, Bu-layev O, Paoli G, Tofanelli S. 2012. The dual origin of tati-speakers from dagestan as written in the genealogy of uniparental variants. Am. J. Hum. Biol. 24:391–399.

Bietti Sestrieri AM. 2010. L'Italia nell'età del bronzo e del ferro. Dalle palafitte a Romo-lo (2200-700 a.C.). Carrocci Editore.

Bini C, Ceccardi S, Lugaresi F, Trane R, Falconi M, Ferri G, Pelotti S. 2008. Polymorphism of mito-chondrial DNA D-loop in Rimini and Valmarecchia areas in the North of Italy. Forensic Sci. Int. Genet. Suppl. Ser. 1:262–263.

Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, Galaxy Team. 2010. Manipulation of FASTQ data with Galaxy. Bioinformatics 26:1783–1785.

Boattini A, Martinez-Cruz B, Sarno S, Harmant C, Useli A, Sanz P, Yang-Yao D, Man-ry J, Ciani G, Luiselli D, et al. 2013. Uniparental markers in Italy reveal a sex-biased ge-netic structure and different historical strata. PLoS One 8:e65441.

Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkošová Z, Powell A, Burger J. 2013. 2000 Years of Parallel Societies in Stone Age Central Europe. Science 342:479-81.

Brisighelli F, Álvarez-Iglesias V, Fondevila M, Blanco-Verea A, Carracedo A, Pascali VL, Capelli C, Salas A. 2012. Uniparental markers of contemporary Italian population-reveals details on its pre-Roman heritage. Caramelli D, editor. PLoS One 7:e50794.

Brizio E. 1895. La necropoli di Novilara, in MALinc5:88-438.

Butler JM, Hill CR. 2012. Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis. Forensic Sci. Rev. 24:15–26.

Capelli C, Brisighelli F, Scarnicci F, Arredi B, Caglia' A, Vetrugno G, Tofanelli S, Ono-fri V, Tagliabracci A, Paoli G, et al. 2007. Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encoun-ter. Mol. Phylogenet. Evol. 44:228–39.

Capocasa M, Anagnostou P, Bachis V, Battaggia C, Bertoncini S, Biondi G, Boattini A, Boschi I, Brisighelli F, Caló CM, et al. 2014. Linguistic, geographic and genetic isola-tion: a collaborative study of Italian populations. J. Anthropol. Sci. = Riv. di Antropol. JASS 92:201–31.

Caramelli D, 2009. Antropologia molecolare. Manuale Base. Firenze University Press.

Caramelli D, Lalueza-Fox C, Vernesi C, Lari M, Casoli A, Mallegni F, Chiarelli B, Dupanloup I, Bertranpetit J, Barbujani G, et al. 2003. Evidence for a genetic discontinu-ity between Neandertals and 24,000-year-old anatomically modern Europeans. Proc. Natl. Acad. Sci. 100:6593–6597.

Caramelli D, Vernesi C, Sanna S, Sampietro L, Lari M, Castrì L, Vona G, Floris R, Francalacci P, Tykot R, et al. 2007. Genetic variation in prehistoric Sardinia. Hum. Ge-net. 122:327–36.

Carfagna B. 2016. I Piceni. Concordanze tra fonti antiche, genetica e archeologia sull'identità e diffusione di un popolo dell'Italia preromana. Capponi Editore.

Chirassi C. 2008. Simbolico Piceno: Un popolo tra mito e storia. In Luni M and Sconocchia S (eds), I Piceni e la Loro Riscoperta tra Settecento e Novecento. Atti del Convegno Internazionale, Urbino 2000. Urbino:Quattroventi, 353-80.

Colonna G, Tagliamonte G. 1999. I popoli del medio adriatico e le tradizioni antiche sulle loro origini. In I Piceni, Popolo d'Europa.Catalogo della Mostra, 2000. Roma: De Luca, 10-3.

Cooper A, Poinar HN. 2000. Ancient DNA: do it right or not at all. Science 289:1139.

Csősz A, Szécsényi-Nagy A, Csákyová V, Langó P, Bódis V, Köhler K, Tömöry G, Nagy M, Mende BG. 2016. Maternal Genetic Ancestry and Legacy of 10th Century AD Hungarians. Sci. Rep. 6:33446.

Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga J-L, et al. 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. 110:15758–15763.

Dabney J, Meyer M, Paabo S. 2013b. Ancient DNA Damage. Cold Spring Harb. Per-spect. Biol. 5:a012567–a012567.

Dalla Fina GM. 2015 (a cura di). La delimitazione dello spazio funerario in Italia dalla Protostoria all'età arcaica. Recinti, circoli, tumuli. Atti del XXII Convegno Internazionale di Studi sulla Storia e l'Archeologia dell'Etruria. Annali della Fondazione per il Museo «Claudio Faina». 22.

De Fanti S, Vianello D, Giuliani C, Quagliariello A, Cherubini A, Sevini F, Iaquilano N, Franceschi C, Sazzini M, Luiselli D. 2016. Massive parallel sequencing of human whole mitochondrial genomes with Ion Torrent technology: an optimized workflow for Anthropological and Population Genetics studies. Mitochondrial DNA Part A:1–8.

Delpino C, Finocchi S, Postrioti G. 2016. Necropoli del Piceno. Dati acquisiti e prospet-tive di ricerca, in Baldini G., Giroldini P.(eds.), "Dalla Val D'Elsa al Conero. Ricerche di archeologia e topografia storica in memoria di Giuliano De Marinis". Atti del Convegno Internazionale di Studi Colle di Val d'Elsa - San Gimignano – Poggibonsi 27-29 no-vembre 2015, All'Insegna del Giglio:287-305.

Delpino C, in press, Infant and child burials in the Picene necropolis of Novilara (Pesaro): 2012-2013 excavations, in Tabolli J (ed), "From invisible to visible. New data and methods for the archaeology of infant and child burials in pre-Roman Italy", Peer reviewed proceedings of the international conference, Trinity College Dublin (24-25 April 2017), in Studies in Mediterranean Archaeology (SIMA).

Der Sarkissian C. 2011. Mitochondrial DNA in ancient human populations of Europe. Doctoral Thesis. https://digital.library.adelaide.edu.au/dspace/handle/2440/74221.

Di Gaetano C, Voglino F, Guarrera S, Fiorito G, Rosa F, Di Blasio AM, Manzini P, Dianzani I, Betti M, Cusi D, et al. 2012. An overview of the genetic structure within the Italian population from genome-wide data. PLoS One 7:e43759.

Di Rienzo A, Wilson AC. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc. Natl. Acad. Sci. U. S. A. 88:1597–601.

Egeland T, Mostad PF, Mevâg B, Stenersen M. 2000. Beyond traditional paternity and identification cases. Selecting the most probable pedigree. Forensic Sci. Int. 110:47–59.

Esser K-H, Marx WH, Lisowsky T. 2006. DNA decontamination: DNA-ExitusPlus in comparison with conventional reagents. Nat. Methods 3.

Excoffier L, Lischer H. 2010. Arlequin suite ver 3.5: a new series of programs to per-form population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567.

Falchi A, Giovannoni L, Calo CM, Piras IS, Moral P, Paoli G, Vona G, Varesi L. 2006. Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. J. Hum. Genet. 51:9–14.

Fiorito G, Di Gaetano C, Guarrera S, Rosa F, Feldman MW, Piazza A, Matullo G. 2016. The Italian genome reflects the history of Europe and the Mediterranean basin. Eur. J. Hum. Genet. 24:1056–1062.

Fortea J, de la Rasilla M, García-Tabernero A, Gigli E, Rosas A, Lalueza-Fox C. 2008. Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain). J Hum Evol 55:353–357.

Fulton TL. 2012. Setting Up an Ancient DNA Laboratory. In: Methods in molecular bi-ology (Clifton, N.J.). Vol. 840. p. 1–11.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. 5:5257.

Ghirotto S, Tassi F, Fumagalli E, Colonna V, Sandionigi A, Lari M, Vai S, Petiti E, Cor-ti G, Rizzi E, et al. 2013. Origins and Evolution of the Etruscans' mtDNA. Hawks J, ed-itor. PLoS One 8:e55519.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blanken-berg D, Albert I, Taylor J, et al. 2005. Galaxy: A platform for interactive large-scale ge-nome analysis. Genome Res. 15:1451–1455.

Goecks J, Nekrutenko A, Taylor J, Galaxy Team T. 2010. Galaxy: a comprehensive ap-proach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 11:R86.

Guimaraes S, Ghirotto S, Benazzo A, Milani L, Lari M, Pilli E, Pecchioli E, Mallegni F, Lippi B, Bertoldi F, et al. 2009. Genealogical discontinuities among Etruscan, Medieval, and contemporary Tuscans. Mol. Biol. Evol. 26:2157–66.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analy-sis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R. 1994. Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am. J. Hum. Genet. 55:175–89.

Hansen HB, Damgaard PB, Margaryan A, Stenderup J, Lynnerup N, Willerslev E, Allentoft ME. 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. Caramelli D, editor. PLoS One 12:e0170940.

Hennessy LK, Mehendale N, Chear K, Jovanovich S, Williams S, Park C, Gangano S. 2014. Developmental validation of the GlobalFiler(®) express kit, a 24-marker STR assay, on the RapidHIT(®) System. Forensic Sci. Int. Genet. 13:247–58.

Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, et al. 2002. Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups. Am. J. Hum. Genet. 70:1152–1171.

Hervella M, Rotea M, Izagirre N, Constantinescu M, Alonso S, Ioana M, Lazăr C, Ridiche F, Soficaru AD, Netea MG, et al. 2015. Ancient DNA from South-East Europe Reveals Different Events during Early and Middle Neolithic Influencing the European Genetic Heritage. Pereira LMSM, editor. PLoS One 10:e0128810.

Hughey JR, Paschou P, Drineas P, Mastropaolo D, Lotakis DM, Navas PA, Michalodimitrakis M, Sta-matoyannopoulos JA, Stamatoyannopoulos G, Broodbank C, et al. 2013. A European population in Minoan Bronze Age Crete. Nat. Commun. 4:1861.

Jin L, Nei M. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. Mol. Biol. Evol. 7:82–102.

Juras A, Dabert M, Kushniarevich A, Malmström H, Raghavan M, Kosicki JZ, Metspalu E, Willerslev E, Pi-ontek J. 2014. Ancient DNA reveals matrilineal continuity in present-day Poland over the last two millennia. PLoS One 9:e110839.

Juras A, Krzewińska M, Nikitin AG, Ehler E, Chyleński M, Łukasik S, Krenz-Niedbała M, Sinika V, Piontek J, Ivanova S, et al. 2017. Diverse origin of mitochondrial lineages in Iron Age Black Sea Scythians. Sci. Rep. 7:43950.

Kefi R, Hechmi M, Naouali C, Jmel H, Hsouna S, Bouzaid E, Abdelhak S, Beraud-Colomb E, Stevanovitch A. 2016 Dec 30. On the origin of Iberomaurusians: new data based on ancient mitochondrial DNA and phylogenetic analysis of Afalou and Taforalt populations. Mitochondrial DNA Part A:1–11.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitu-tions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–20.

Kling D, Tillmar AO, Egeland T. 2014. Familias 3 - Extensions and new functionality. Forensic Sci. Int. Genet. 13:121–7.

Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F. 2011. HaploGrep: a fast and reliable algorithm for automatic classifica-tion of mitochondrial DNA haplogroups. Hum. Mutat. 32:25–32.

Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith EA. 2012. Setting the stage? Building and working in an ancient DNA laboratory. Ann. Anat. - Anat. Anzeiger 194:3–6.

Knapp M, Lalueza-Fox C, Hofreiter M. 2015. Re-inventing ancient human DNA. Inves-tig. Genet. 6:4.

Knipper C, Meyer C, Jacobi F, Roth C, Fecher M, Stephan E, Schatz K, Hansen L, Posluschny A, Höppner B, et al. 2014. Social differentiation and land use at an Early Iron Age "princely seat": bioarchaeological investigations at the Glauberg (Germany). J. Archaeol. Sci. 41:818–835.

Kruskal JB. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29:1–27.

Krzewińska M, Bjørnstad G, Skoglund P, Olason PI, Bill J, Götherström A, Hagelberg E. 2015. Mitochondrial DNA variation in the Viking age population of Norway. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 370:20130384.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Anal-ysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33:1870–1874.

LaRue BL, Ge J, King JL, Budowle B. 2012. A validation study of the Qiagen Investigator DIPplex® kit; an INDEL-based assay for human identification. Int. J. Legal Med. 126:533–40.

Le Roy M, Rivollat M, Mendisco F, Pemonge M-H, Coutelier C, Couture C, Tillier A, Rottier S, Deguilloux M-F. 2016. Distinct ancestries for similar funerary practices? A GIS analysis comparing funerary, osteological and aDNA data from the Middle Neo-lithic necropolis Gurgy "Les Noisats" (Yonne, France). J. Archaeol. Sci. 73:45–54.

Lincoln PJ. 1997. DNA recommendations--further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Forensic Sci. Int. 87:181–4.

Lollini DG, 1976. La civiltà picena, in Popoli e civiltà dell'Italia antica V, Roma:107-195.

Lorkiewicz W, Płoszaj T, Jędrychowska-Dańska K, Żądzińska E, Strapagiel D, Haduch E, Szczepanek A, Grygiel R, Witas HW. 2015. Between the Baltic and Danubian Worlds: the genetic affinities of a Middle Neolithic population from central Poland. PLoS One 10:e0118316.

Matney T, Algaze G, Dulik MC, Erdal ÖD, Erdal YS, Gokcumen O, Lorenz J, Mergen H. 2012. Understanding Early Bronze Age social structure through mortuary remains: A pilot aDNA study from Titriş Höyük, southeastern Turkey. Int. J. Osteoarchaeol. 22:338–351.

Modi A, Tassi F, Susca RR, Vai S, Rizzi E, Bellis G De, Lugliè C, Gonzalez Fortes G, Lari M, Barbujani G, et al. 2017. Complete mitochondrial sequences from Mesolithic Sardinia. Sci. Rep. 7:42869.

Naso A. 2000. I Piceni. Storia e archeologia delle Marche in epoca preromana. Longa-nesi &C.

Nei M, 1987. Molecular Evolutionary Genetics. Columbia University Press.

Nikitin AG, Potekhina I, Rohland N, Mallick S, Reich D, Lillie M. 2017. Mitochondrial DNA analysis of eneolithic trypillians from Ukraine reveals neolithic farming genetic roots. PLoS One 12:e0172952.

Olivieri A, Sidore C, Achilli A, Angius A, Posth C, Furtwängler A, Brandini S, Capodi-ferro MR, Gandini F, Zoledziewska M, et al. 2017. Mitogenome Diversity in Sardinians: a Genetic Window onto an Island's Past. Mol. Biol. Evol.

Pääbo S, Higuchi RG, Wilson AC. 1989. Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. J. Biol. Chem. 264:9709–12.

Palencia-Madrid L, Cardoso S, Keyser C, López-Quintana JC, Guenaga-Lizasu A, de Pancorbo MM. 2017. Ancient mitochondrial lineages support the prehistoric maternal root of Basques in Northern Iberian Peninsula. Eur. J. Hum. Genet. 25:631–636.

Palencia-Madrid L, de Pancorbo MM. 2015. Next generation sequencing as a suitable analysis for authentication of degraded and ancient DNA. Forensic Sci. Int. Genet. Suppl. Ser. 5:e338–e340.

Pereira R, Phillips C, Alves C, Amorim A, Carracedo Á, Gusmão L. 2009. A new multiplex for human identification using insertion/deletion polymorphisms. Electrophoresis 30:3682–3690.

Pesando F. 2005. L'Italia antica. Culture e forme del popolamento nel I millennio a. C. Roma: Carocci editore. 326.

Piazza A, Cappello N, Olivetti E, Rendine S. 1988. A genetic history of Italy. Ann. Hum. Genet. 52, 203–213.

Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, Gerritsen F, Moiseyev V, Gromov A, Raczky P, et al. 2015. Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLoS One 10:e0129102.

Presciuttini S, Cerri N, Turrina S, Pennato B, Alù M, Asmundo A, Barbaro A, Boschi I, Buscemi L, Caenazzo L, et al. 2006. Validation of a large Italian Database of 15 STR loci. Forensic Sci. Int. 156:266–8.

Previderè C, Grignani P, Alessandrini F, Alù M, Biondo R, Boschi I, Caenazzo L, Car-boni I, Carnevali E, De Stefano F, et al. 2013. The 2011 GeFI collaborative exercise. Concordance study, proficiency testing and Italian population data on the new ENFSI/EDNAP loci D1S1656, D2S441, D10S1248, D12S391, D22S1045. Forensic Sci. Int. Genet. 7:e15-8.

Prüfer K, Stenzel U, Hofreiter M, Pääbo S, Kelso J, Green RE. 2010. Computational challenges in the analysis of ancient DNA. Genome Biol. 11:R47.

Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, Morel N, Fernandez-Jalvo Y, Grange T, Geigl E-M. 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. Proc. Natl. Acad. Sci. U. S. A. 104:739–44.

R Development Core Team 2008. R: A language and environment for statistical compu-ting. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, et al. 2000. Tracing European founder lineages in the Near East-ern mtDNA pool. Am. J. Hum. Genet. 67:1251–76.

Rivollat M, Mendisco F, Pemonge M-H, Safi A, Saint-Marc D, Brémond A, Couture-Veschambre C, Rottier S, Deguilloux M-F. 2015. When the waves of European Neo-lithization met: first paleogenetic evidence from early farmers in the southern Paris Ba-sin. PLoS One 10:e0125521.

Rizzi E, Lari M, Gigli E, De Bellis G, Caramelli D, Magnussen K, Steinmann K, Kapra-nov P, Thompson J, Zazula G, et al. 2012. Ancient DNA studies: new perspectives on old samples. Genet. Sel. Evol. 44:21.

Sampietro ML, Caramelli D, Lao O, Calafell F, Comas D, Lari M, Agustí B, Bertranpetit J, Lalueza-Fox C. 2005. The genetics of the pre-Roman Iberian Peninsula: a mtDNA study of ancient Iberians. Ann. Hum. Genet. 69:535–48.

Sampietro ML, Gilbert MTP, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006. Tracking down Human Contamination in Ancient Human Teeth. Mol. Biol. Evol. 23:1801–1807.

Sarno S, Boattini A, Carta M, Ferri G, Alù M, Yao DY, Ciani G, Pettener D, Luiselli D. 2014. An Ancient Mediterranean Melting Pot: Investigating the Uniparental Genetic Structure and Population History of Sicily and Southern Italy. Caramelli D, editor. PLoS One 9:e96074.

Sarno S, Boattini A, Pagani L, Sazzini M, De Fanti S, Quagliariello A, Gnecchi Ruscone GA, Guichard E, Ciani G, Bortolini E, et al. 2017. Ancient and recent admixture layers in Sicily and Southern Italy trace multiple migration routes along the Mediterranean. Sci. Rep. 7:1984.

Sazzini M, Gnecchi Ruscone GA, Giuliani C, Sarno S, Quagliariello A, De Fanti S, Boattini A, Gentilini D, Fiorito G, Catanoso M, et al. 2016. Complex interplay between neutral and adaptive evolution shaped differential genomic background and disease sus-ceptibility along the Italian peninsula. Sci. Rep. 6:32513.

Simón M, Jordana X, Armentano N, Santos C, Díaz N, Solórzano E, López JB, González-Ruiz M, Malgosa A. 2011. The presence of nuclear families in prehistoric collective burials revisited: The bronze age burial of montanissell cave (Spain) in the light of aDNA. Am. J. Phys. Anthropol. 146:406–413.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele fre-quencies. Genetics 139:457–62.

Stiller M, Knapp M, Stenzel U, Hofreiter M, Meyer M. 2009. Direct multiplex sequenc-ing (DMPS)--a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. Genome Res. 19:1843–1848.

Tagliabracci A, 2009. Introduzione alla genetica forense – Indagini di identificazione personale e di paternità. Springer Biomed.

Tagliamonte G, 1994. I Figli di Marte. Mobilità, Mercenari a Mercenariato Italici in Magna Grecia e Sicilia. Roma, Giorgio Bretschneider.

Tagliamonte G. 1996. I Sanniti. Caudini, Irpini, Pentri, Carricicni, Frentani. Milano:Longhesi.

Tajima F. 1993. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135:599–607.

Tajlma F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105(2):437-60.

Turchi C, Buscemi L, Previderè C, Grignani P, Brandstätter A, Achilli A, Parson W, Tagliabracci A, Group GFI. 2008. Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. Int. J. Legal Med. 122:199–204.

Vai S, Ghirotto S, Pilli E, Tassi F, Lari M, Rizzi E, Matas-Lalueza L, Ramirez O, Lalue-za-Fox C, Achilli A, et al. 2015. Genealogical relationships between early medieval and modern inhabitants of Piedmont. PLoS One 10:e0116801.

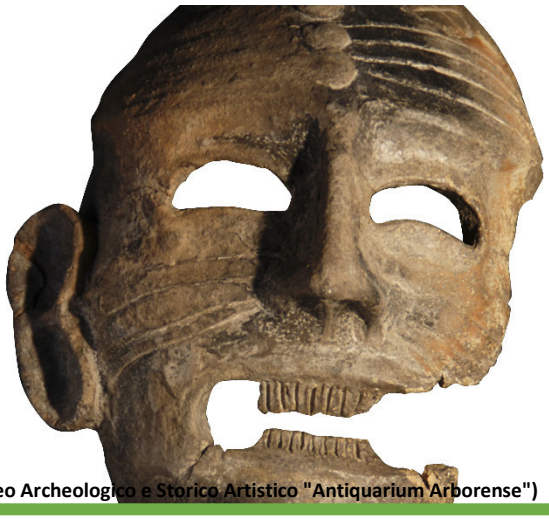
Van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global hu-man mitochondrial DNA variation. Hum. Mutat. 30:E386–E394.

Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castrì L, Casoli A, et al. 2004. The Etruscans: a population-genetic study. Am. J. Hum. Genet. 74:694–704.

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc. Natl. Acad. Sci. U. S. A. 86:9350–4.

Wang DY, Gopinath S, Lagacé RE, Norona W, Hennessy LK, Short ML, Mulero JJ. 2015. Developmental validation of the GlobalFiler® Express PCR Amplification Kit: A 6-dye multiplex assay for the direct amplification of reference samples. Forensic Sci. Int. Genet. 19:148–155.

Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, Unterländer M, Hollfelder N, Potekhina ID, Schier W, Thomas MG, et al. 2014. Direct evidence for positive selection of skin, hair, and eye pigmentation in Eu-ropeans during the last 5,000 y. Proc. Natl. Acad. Sci. U. S. A. 111:4832.



Apotropaic mask from the necropolis of Tharros (6th century BC) (Museo Archeologico e Storico Artistico "Antiquarium Arborense")

CASE STUDY II

Deciphering the identity and settlement of the "Phoenician-Punic" civilization: a first genetic study on Tharros (OR) and Lilybaeum (TP) sites

Abstract: The timing and modalities concerning the identity and expansion of the "Phoenician" civilization and the formation and diffusion of the "Punic" culture - linked to the Carthage cultural and territorial expansion - represent in the Phoenician-Punic studies a vexed question. In order to contribute to the reconstruction of the "Phoenician-Punic" settlement in the central-western Mediterranean area, a research project has been started in the Tharros southern necropolis (OR, Sardinia) and Lilibaeum site (TP, Sicily) based on a multidisciplinary approach that combines the contributions of archaeological and genetic investigations *in primis*. In the present study, conducted on a first selection of bone samples, classical methods of mitochondrial DNA analysis (HVS-I and SNPs of the coding region) have been combined with new generation techniques (NGS) to obtain ancient whole genomes. This research, therefore, provides a pioneer survey in the Phoenician-Punic context, to define the target population and expand the knowledge on migration flows and the relationship between ancient and present-day populations of the Mediterranean area, to trace the ethnic origin, and to understand whether will be maintained a genetic continuity with those who nowadays still live in the same territories.

Keywords: Iron Age, Phoenician-Punics, Tharros, Lilibaeum, ancient DNA, mitochondrial DNA, HVS-I region, Next-Generation Sequencing, whole genome, kinship relationship, population genetic.

5.1 Introduction

5.1.1. The origin of the Phoenician-Punic civilization: historical and cultural background

Usually, we referred as 'Phoenician civilization' to the culture that developed in the eastern coasts of the Mediterranean, that coincide with the area of present-day Lebanon, from the end of the 2nd Millennium BC to the Hellenistic period (Bondì, 2009). The existence of a *caesura*

between the ‘Canaanite civilization’ of the Bronze Age and the ‘Phoenician civilization’ developed at the beginning of the Iron Age has been investigated based on archaeological (Aubet, 2009) and linguistic evidence (Xella, 2014). Some researchers also proposed that the geographical boundaries of Phoenician influence extended during the 1st Millennium BC from Mount Cassius in Syria to Dor in Israel (e.g. Gras et al., 1995; Lipinski, 1995).

Defining the geographical and chronological boundaries of Phoenician and Punic civilization remains a complex issue. In fact, the same terminology used to define the ‘Phoenicians’ is controversial, and to define it is therefore important because of the ethnic, linguistic, geographical, and cultural implications of the terms Phoenician and Punic (Aubet, 2009). The term ‘Phoinike’ appears for the first time in Homer and Hesiod, between 9th and 7th century BC, and was used from Greek-speaking people to indicate the territory that extended from Arwad to the North up to the Carmel mountain to the South (Moscati, 1988; Mazza, 1995). The meaning of the ethnonym used by the Greeks for the ‘Phoinikes’ foreign people is not clear, and it has been related to the term ‘phoinós’, red colour, referring to the skin complexion of these people. Alternatively, the term has been read as an allusion to the purple industry, for which the Phoenician cities were famous in the times of Homer, or as a reference to the eponymous hero Phoinix, inventor of the technique of purple extraction according to Pliny the Elder (Bonnet, 1983; Aubet, 2009). According to another interpretation the term was born in the Mycenaean area of Crete in the Late Bronze Age, with the meaning of ‘red-skinned people’ based on the assonance of the term ‘po-ni-ki-jo’ with the word ‘fenkhou’, this last used in the Egyptian language to indicate an Asian population settled in an area next to Egypt (Bondì, 2009).

Nevertheless, no one knew how the Phoenicians defined themselves: it has been suggested that they may have used the Semitic term of ‘Canaanite’ (Aubet, 2009; Bondì 2009). Neither ‘Canaan’ nor ‘Canaanites’, however, refer to ethnic and national realities, but indicate a very large geographical entity in the Syro-Palestinian context which, during the Late Bronze Age, constituted one of the three districts of the territory subject to Egyptian control, corresponding to the present-day Palestine (Xella, 2014). Most likely, the Phoenicians referred to themselves through the name of their city of origin (i.e. Sidonians, Tyrians, Giblites) as recorded in Assyrian annals and in Homer’s poems (Vella, 1998; Xella, 2014). The Roman authors used the terms ‘Poenus’ and ‘Phoenix’, transcribed from the Greek, to indicate the Phoenicians in general and specifically the Carthaginians. The noun ‘Poenus’, as well as the adjectives poenicus and punicus, were used as synonyms to refer to the North African Phoenicians. In

modern historiography, however, the terms 'Phoenician' and 'Punic' assumed ever more marked geographical and chronological connotations. In the scientific literature, the term 'Phoenicians' is conventionally used to indicate the Phoenicians of both Eastern and Western area, in the historical period preceding the political and military hegemony reached by Carthage at the end of the 6th century BC. Conversely, the term 'Punic' is used when referring to Phoenicians of the Western area, between the 6th century BC until the fall of Carthage in 146 BC. While the term 'Carthaginian' means people, events, and culture concerning the city of Carthage (Aubet, 2009; Bondì, 2009).

The assignment to these historiographic labels of too marked and rigid spatial, temporal and cultural values shows sometimes some difficulties deriving from attributing complex phenomena to too rigid schemes of interpretation (see i.e. the so-called 'Phoenician' inhumation in the necropolis of Tharros and Othoca, see Del Vais and Fariselli, 2010, 2012; Del Vais and Usai, 2013; Fariselli, 2017).

5.1.2. The Phoenicians and the West: the colonization

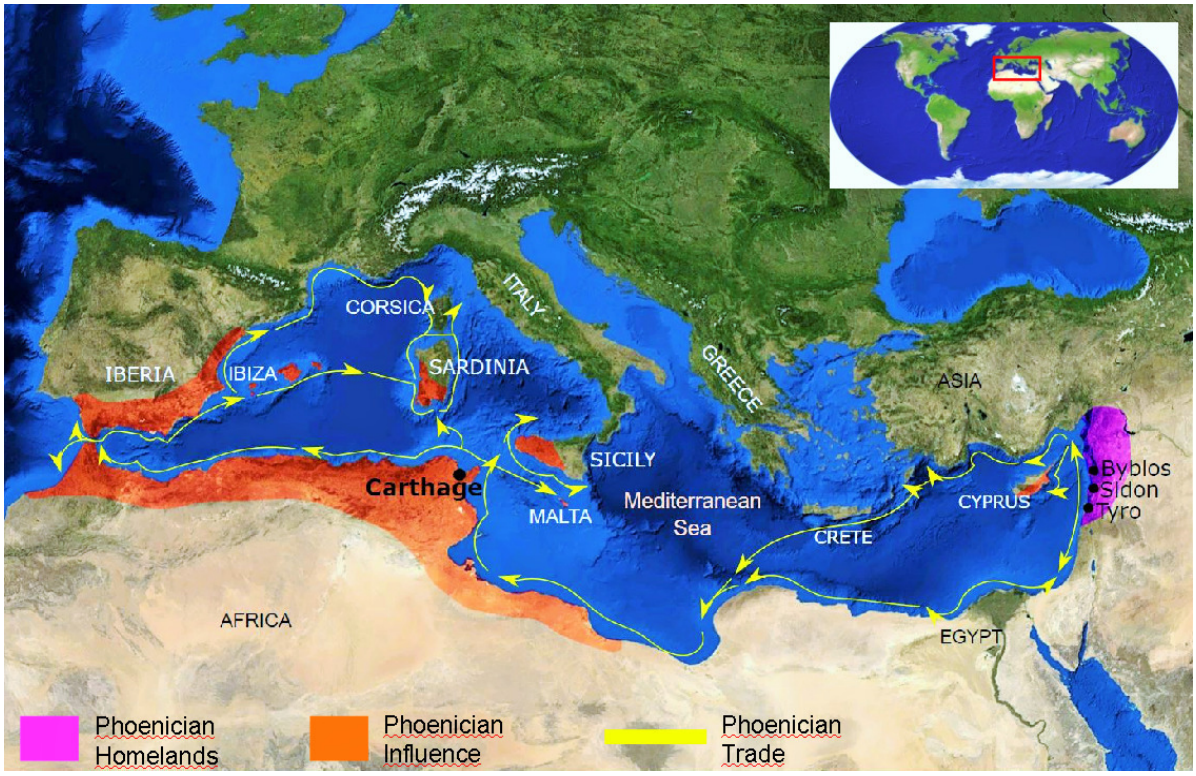


Figure 5.1.2.1 | Phoenician trade routes during the 6th century BC.

The Phoenician homeland, which never constituted a unitary political structure, was a narrow coastal plain closed by the Lebanese mountain ranges and divided into several areas by the presence of several rivers. The main autonomous city-states, located on the mainland coastline or on tiny offshore islands facing the coast, were Aradus, Tripolis, Byblos, Berytus, Sidon and Tyre.

In the 1st Millennium BC, Phoenician influence spread North and South of the coastal plain described, and some modern authors would see the Phoenician territory extending beyond these confines to include an area from Mount Cassius in Syria to Dor in Israel (e.g. Gras et al., 1995; Lipinski, 1995).

Phoenician colonization was preceded by a phase (in literature improperly referred to as 'Pre-colonization') of episodic frequentation limited to some specific Mediterranean areas, above all Spain and Sardinia, with the goal of starting contacts with business partners. Phoenician colonization is a phenomenon consisting in several immigration/colonial waves that, from the end of the 9th and the beginning of the 8th century BC to the end of the 7th century BC, involved the entire Mediterranean basin and the Atlantic coasts of the Iberian Peninsula and Africa. It has long been considered as an effect of the oppression exercised by the Assyrians on the cities of Phoenicia but more recently the whole phenomenon has been reinterpreted as a result of the development in the social, political and economic situation both within the same Phoenician cities and in the international context.

The city of Tyre played a fundamental role in the colonial process and among its numerous foundations there are Kition in Cyprus and, above all Carthage in North Africa, which was founded at the end of the 9th century BC according to literary sources (Timeo of Tauromenio apud Dionigi of Alicarnasso, Ant. rom. I, 74, 1; Flavio Giuseppe, Cont. Ap., I, 125; Giustino, XVIII, 6, 1-9).

Following the Mediterranean routes (Figure 5.1.2.1), the Phoenicians reached and founded colonies in:

- Malta;
- Sicily: Mozia, Lilybaeum (Figure 5.1.3.1), Palermo and Solunto;
- North Africa: besides Carthage, the settlements of Leptis Magna, Sabratha, Sousse, Utica, Mersa Madakh, Rachgoun, Lixus, Sala and Mogador;
- Sardinia: Cuccureddus of Villasimius, Cagliari, Nora, Bitia, Sulky, Monte Sirai, Neapolis, Othoca, Tharros (Figure 5.1.3.1), Bosa, Olbia;

- Balearics (Sa Caleta in Ibiza);
- Iberian Peninsula (La Fonteta, Villaricos, Almuñecar, Chorreas, Morro de Mezquitilla, Toscanos, Cerro del Villar, Cadiz, Castillo de Doña blanca, Abul, Santa Olaia). (see Bondì, 2009).

5.1.3. Archaeological context of the samples: from Sardinia to Sicily

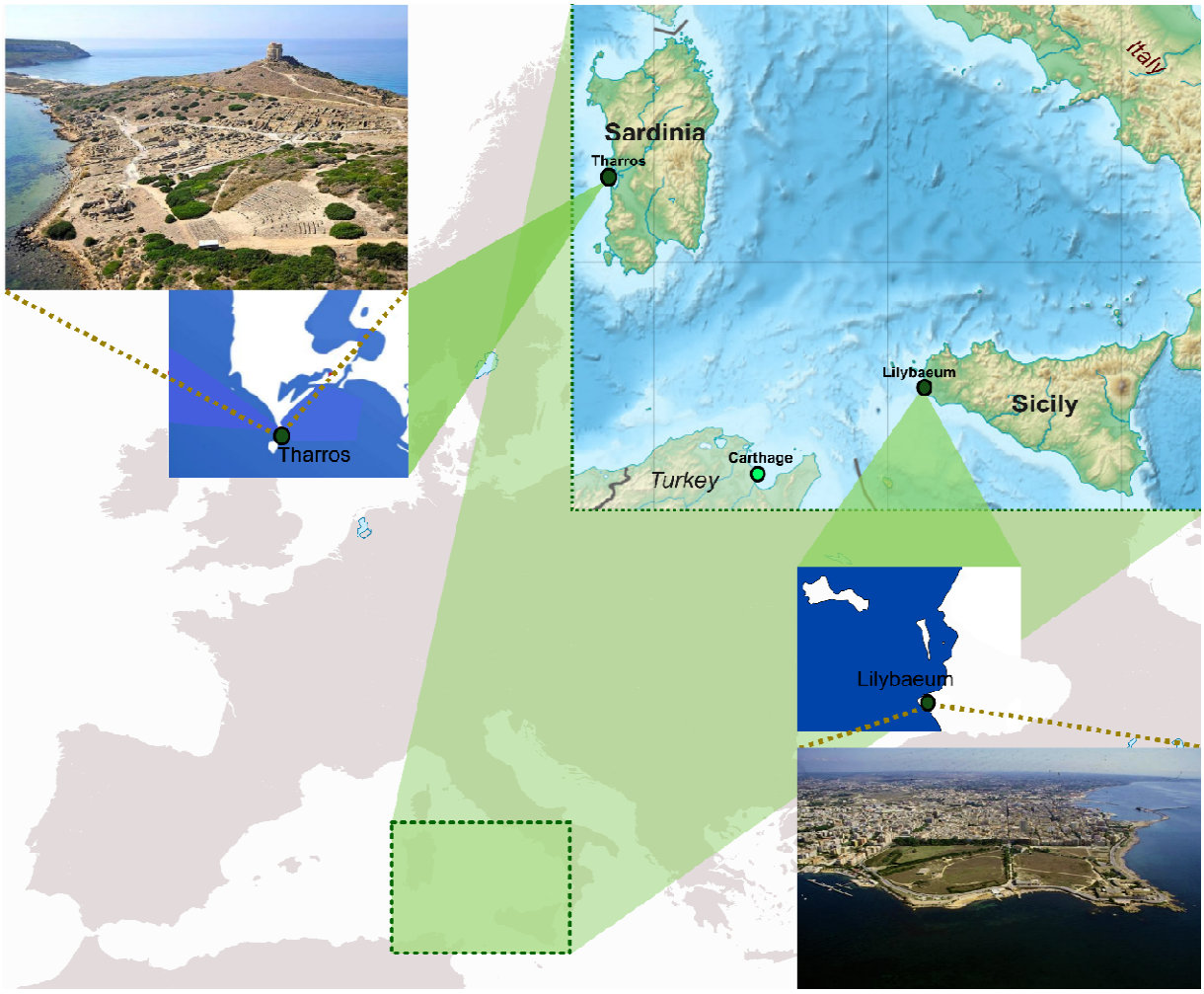


Figure 5.1.3.1 | Geographic location of the necropolis analysed: a) aerial photo of Tharros site (OR, Sardinia); b) aerial photo of Lilybaeum site (TP, Sicily).

5.1.3.1. The ancient site of Tharros (OR, Sardinia)

The ancient Phoenician and Punic colony of Tharros was established in the 7th century BC in the Sinis Peninsula, in an area already occupied by Nuragic populations (Depalmas and Melis, 2010; Usai, 2014). The exact location of the archaic settlement of Tharros is currently unknown and the oldest archaeological records come from the necropolis and from the tophet,

the Punic sanctuary of children. Recent surveys conducted in the lagoon of Mistras, used as a naturally sheltered harbour, have documented the presence of an intensive trading area connected to the city of Tharros since almost the 7th century BC (Pascucci et al., 2018). During the 6th century BC, Sardinia entered definitively in the sphere of influence of Carthage and Tharros probably with a leading role in the administrative and economic management of the Island, becoming itself 'Carthage of Sardinia' (Acquaro, 1995). According to the importance assumed during the Punic Age, Tharros underwent various urban interventions such as the installation of a metallurgical handicraft next to the tophet in the northern district of the city and the construction or restoration of public buildings, like the so-called 'monumental temple' (Floris, 2015) and the city walls (Acquaro and Mezzolani, 1996). After the Roman conquest of Sardinia in 238 BC, the residential quarters of the Punic city were preserved substantially unchanged (Marano, 2017) but urban layout of Tharros underwent to numerous transformations: the city walls and temples were renewed, streets were paved with basaltic slabs, large buildings typical of Romanization – such as the baths – were built (Acquaro and Mezzolani, 1996).

After a progressive decline in Late Antiquity and Early Middle Ages (5th–6th century AD), Tharros and the Sinis region were abandoned in the Middle Ages due to the raids of Saracens along the coastline (Spanu, 1998).

The ancient site of Tharros is unfortunately well known since the 17th century when the richness of the burial goods of the cemeteries of the city attracted the attention of treasure hunters. Even if some explorations of the cemeteries started in the 19th century, the first scientific survey was carried out in the so-called 'southern necropolis' of Capo San Marco by G. Spano around 1850s (Spano, 1851). The following years were characterized by a real 'gold rush' and by the wild looting of the necropolis, started by the excavations of Lord Vernon, and by the official explorations between 1853 and 1856 of G. Cara, the future Chief of the Royal Museum of Cagliari, which were followed by the illegal sale of most of the finds, only partially converged in the British Museum in London. At the end of the 19th century, between 1885 and 1886, new excavations were directed by the Royal Inspector of Antiquities F. Nissardi, who provided an accurate topographic plan from the northern to the southern necropolis (Del Vais, 2006).

The urban area of Tharros was investigated between 1956 and 1964 by the Superintendent G. Pesce, who discovered most of the Punic-Roman city visible nowadays (Pesce, 1966; Marano, 2014), and the tophet, located in the Northern district of the city, on the East side of

Su Murru Mannu hill (Fariselli, 2015). In the second half of the 1960s the explorations were directed by F. Barreca, the future Superintendent, who excavated the so-called 'temple of Demeter' (Floris, 2016), and the fortifications of Su Murru Mannu (Acquaro and Mezzolani, 1996). From 1974 to 1996 the Northern area of the city, and specifically the tophet (Fariselli 2015) and the nearby metallurgical handicraft area (Del Vais, 2015), were investigated by the joint mission of the 'Institute for the Phoenician and Punic Civilization of the National Research Centre (CNR)', the Archaeological Superintendence of Cagliari and Oristano, and the University of Bologna (Acquaro et al., 1974-2000). This effort was directed in the first year by A. Ciasca, subsequently by E. Acquaro, and finally by M.T. Francisi.

The resumption of systematic investigations in the 'southern necropolis' took place between 2001 and 2004 by a joint mission of the Archaeological Superintendence and the University of Bologna in collaboration with the University of Cagliari (Acquaro et al., 2006; Fariselli, 2008). The so-called 'northern necropolis', located in correspondence of the modern village of San Giovanni di Sinis, was the object of the excavations of the University of Cagliari directed by C. Del Vais from 2009 to 2013 (Del Vais and Fariselli, 2010a, 2010b; Del Vais, 2013; Fariselli, 2013a). A new ongoing research project, which resumes the archaeological excavations in the 'southern necropolis' under the direction of A.C. Fariselli of the University of Bologna (Fariselli, 2014; Secci, 2016; Fariselli, 2017), has resulted in the discovery of numerous Punic burial chambers and pits carved in the rocky bank and the retrieval of a variety of burial goods.

5.1.3.1.1. The cemeteries of Tharros: an overview

As described above, the Phoenician and Punic city of Tharros was served by two cemeteries: a) the 'northern necropolis', situated inside the modern village of San Giovanni di Sinis (Del Vais and Fariselli, 2010a, 2010b; Del Vais, 2013; Fariselli, 2013a), and b) the 'southern necropolis', that is located on Capo San Marco area (Acquaro et al., 2006; Fariselli, 2008; Fariselli, 2014; Secci, 2016; Fariselli, 2017). Both cemeteries were used for the same chronological range, and archaeological investigations carried out to date have documented the presence of similar tomb types and burial practices.

In the first centuries since the foundation of the colony, in the 'Archaic age' (7th -6th centuries BC), the most common funerary practice was incineration, and primary and secondary depositions are both documented. In primary depositions, the dead's body laid upon the funeral pyre situated inside of the elliptical-shaped pit. After the combustion, the

funerary goods were set at the bottom of the pit beside the bones laid in anatomical position and the burnt wood. The pit was then filled up and closed with stone slabs (Del Vais and Fariselli, 2010; Del Vais and Fariselli, 2012). In secondary depositions, the dead was cremated in an *ustrinum* (Del Vais and Fariselli, 2012) and the bones collected and placed in sub-circular and small elliptic pits (Del Vais and Fariselli, 2010; Del Vais and Fariselli, 2012). Only in the 'southern necropolis' the grave seldom consisted of a lithic cyst (Del Vais, 2006). The skeletal remains were set on the bottom of the pit, more rarely in ceramic urns. The human remains were accompanied by personal funeral equipment, such as jewellery, scarab-shaped seals, amulets, bronze weapons (Fariselli, 2013a), and in an almost standardized set of ritual vessels (Del Vais and Fariselli, 2010).

During the 'Punic age', in a period that some authors make it correspond with the Carthaginian military conquest of Sardinia, in the second half of the 6th century BC, according to the literary texts (although the degree of reliability of these sources is debated, see i.e. Krings, 2000), cremation ritual was replaced by inhumation (Bartoloni, 1981). However, recent surveys have documented the use of the inhumation rite in archaic tombs, which can be chronologically placed between the end of the 7th and the beginning of the 6th century BC (Del Vais and Fariselli, 2010; Del Vais and Fariselli, 2012). In this period, new tomb types were introduced, such as rock-cut parallelepiped pits, variable in depth, and hypogea graves. The latter type was composed of a dromos, a rectangular entrance module with steps either continuous or carved along the sidewall, and a square funeral chamber (Fariselli, 2006; Del Vais and Fariselli, 2010).

In some cases, both in the hypogea and in the monumental-sized pit tombs, niches were dug in the rocky walls and little wells were made on the rocky floor (Secci, 2016). In the northern as well as in the southern necropolis some pits and hypogea show a relief decoration. The motifs depicted are aniconic religious symbols: betyls are represented more frequently, but the lozenge and the so-called 'bottle idol' are also attested (Fariselli, 2006; Del Vais and Fariselli, 2010; Del Vais and Fariselli, 2012). The chambers were regularly closed with a slab of stone and the dromos filled with residual sandstone flakes from the excavation of the tomb (Fariselli, 2006). The pit tombs were closed by a lithic cover housed within a recess cut at the top of the rock faces. The cover usually consisted of several stone slabs placed side by side (Fariselli, 2006) or, more rarely, of a monolith with a double-sloping top and equipped with a small altar (Del Vais, 2013). The deceased were laid on the rocky floor of the tomb or inside a wooden sarcophagi or coffin, documented by the discovery of nails and split pins (Fariselli,

2006), in supine position with their arms along the body or folded. The bodies were accompanied by ceramic vessels and objects of personal adornments, such as amulets, scarabs and golden or silver jewels (Fariselli et al., 2012; Fariselli, 2013a). The funerary landscape was characterized by the presence of cippi, stelae and small altars embedded in the ground to signal the burials (Del Vais, 2011; Del Vais, 2013).

In Roman Age, the cemeteries had a higher extent than in Punic age and occupied even the moat of the fortifications in the Northern slopes of Su Murru Mannu hill, the Southern slope of the San Giovanni hill, and the area around the early Christian church of San Giovanni (Acquaro and Mezzolani, 1996). In this period are documented the funerary practices of cremation and inhumation and, in addition, of an occasional re-utilization of Punic tombs, new tomb types are known, such as 'Capuchin' style burials, earth pits, mausoleums, 'cupa' type tombs and lythic *sarcophagi* (Zucca, 1984).

5.1.3.2. The ancient site of Lilybaeum (TP, Sicily)

In 397 BC the Phoenician city of Motya, situated on the island of San Pantaleo at the centre of a large lagoon, known today as 'Stagnone', was invaded and destroyed by the Syracusan tyrant Dionysus I (Diod. Sic. 22.10.4). The survivors founded a town on the mainland nearby, on the promontory of Capo Boeo, that they called Lilybaeum, that developed into the most important military stronghold in Punic Sicily. The city's name probably comes from the Greek word 'Lilýbaion', which means "one that guards Libya" or, according to Diodorus of Sicily (Diod. Sic. 12.54.4), it could derive from the homonymous water well now absorbed by the Church of San Giovanni at Boeo (Caruso, 2008).

The city of Lilybaeum (Di Stefano, 1993; Caruso, 2000) covers a large square area partly bordered by the sea; the sides facing the mainland were defended by a deep moat, ranging between 5.6 to 7 meters, and strong towered ramparts. A vast necropolis ran along the North-East wall, beyond the moat (Becthold, 1999). Diodorus describes the secure nature of the settlement in that it was surrounded by the sea with Carthaginian fortifications on the inland side (Diod. Sic. 22.10.5). Thanks to its imposing fortifications and to the natural canal of dunes and cliffs (that linked the harbour to the 'Stagnone' lagoon making access difficult because of shallow waters), Lilybaeum was the only city within the Carthaginian *epikrateia* not to have fallen to Pyrrhus, king of Epirus, in 278 BC, and Hanno is claimed to have gathered his forces here in 264 BC before advancing to Solus (Diod. Sic. 22.10.4; and 23.1.1). According to the historian Polybius, Lilybaeum was the stronghold that enabled the

Carthaginians to maintain their control of western Sicily during the First Punic War against Roman (Polyb. 1.41.4-6, and 1.42.8-9). Despite 10 years of siege and strict naval blockade, the town resisted Roman conquest and the Punic troops were evacuated only after the peace treaty (241 AD) that put an end to the war (Diod. Sic. 24.14.1; Polyb. 1.41-42, and 1.44-45). Lilybaeum was incorporated into the early province of Sicily, becoming the seat of the first praetor (Livy 22.31.6, 23.31.2, and 31.29.8). After that, Carthage tried to regain Lilybaeum during the Second Punic War, albeit without success, and the settlement remained an important harbour for the Romans, especially when they focused their attack on Carthage (Livy 21.49.2-7, 21.50.10-21.51.1, 22.56.7, 23.21.2, 25.31.12-14, 27.5.9, 28.4.5-7). Lilybaeum continued to act as capital until 211 BC when this was moved to Syracuse, but it is also written that the province retained the position of two quaestors, one in Syracuse with *imperium provinciae* over Hieron's former kingdom, and one in Lilybaeum with a similar command over the *vetus provincia*. In 74-75 AD, Cicero refers to Lilybaeum as a most splendid city, *splendissima civitas*, with a community of Roman citizens (Cic. Verr. 2.5.4.10, 2.5.54.140). The town's economy further developed during the Roman Empire because of its strategic position along the commercial maritime routes from Northern Africa to Rome; the ruins of several luxurious private dwellings, with a wealth of thermal baths and polychrome mosaics, brought to light by during excavations in Lilibeo, date back to that period (Strabo 6.2.5).

With Rome's decline, Genseric's Vandals attacked and virtually destroyed Marsala in 440 AD. The site maintained its role as a crucial maritime port also under Arab and Norman rule: travellers of that period often referred to Lilybaeum and described the town. In fact, it was during the years of the Arab domination that the city was named Marsala, from the Arab 'Mars el Allah', or 'God's Harbour.

5.1.3.2.1 The necropolis of Lilybaeum: an overview

The Punic-Roman necropolis of Lilybaeum (Becthold, 1999) is located beyond the defensive moat, along with the North-Western and North-Eastern wall of the city, and was used from the foundation of the city until the Imperial Age (4th century BC- 2nd century AD).

The existence of Lilybaeum cemetery was known since the end of the 18th century, thanks to the discovery of a crater - cratere Grignani – that was used for a Punic tomb as a cinerary urn, and thanks to the Schubring's studies and to the work by Francesco Struppa, which

collected a series of objects casually found in the cemetery area (Giglio and Canzonieri, 2009). The systematic exploration of the Punic necropolis was carried out by Salinas in 1894 thanks to a series of archaeological excavations in a Western area of the current Whitaker Avenue. These findings provided important results, allowing to outline the presence of some recurring types of burials since the earliest phases of the necropolis (Giglio and Canzonieri, 2009).

In 1902, Salinas explored some areas in the locality of 'Pozzallo alla Salinella', that allowed him to discover a series of funerary *aediculae*, which were obtained by carving the tender local rock. Further investigations were carried out in 1919 by B. Peace: he collected a series of information about some Punic graves explored in the area of 'Bastione S. Francesco' (Marsala) (Giglio and Canzonieri, 2009). A systematic research began only after the Second World War, following the expansion of the reconstruction of Marsala, around the areas between 'Via del Fante' and 'Corso Gramsci' city roads. The scientific research conducted by A.M. Bisi brought to light the 'oldest part' of the Lilybaeum necropolis, located in the area bordering Cape Boeo *insula* (Giglio and Canzonieri, 2009). Afterwards, in 1970, C.A. Di Stefano, on behalf of the 'Soprintendenza Archeologica della Sicilia Occidentale', contributed to the knowledge of the Punic-Roman necropolis of Lilybaeum through the discovery of several burials. In 1987, the ongoing construction activity in the area described above required a series of new interventions in 'via Cattaneo' and 'Corso Gramsci' areas, where hypogenic contexts were discovered. Another important piece was added thanks to 'via Berta' area excavations in 1991, aimed at completing the excavations started in 1985 (Giglio and Canzonieri, 2009).

The typical rock-cut Punic graves either have a vertical shaft leading into one or more funerary chambers (type I) - probably destined to families of high social status - or mostly consist of a rectangular cist (type II), inside which the body of the deceased was deposited on a wooden support. The archaeological evidence revealed that these tombs were utilized for inhumation burials: human remains have been found in most cases not in anatomical connection and in extreme fragmentation conditions. Although rare, the ritual of the incineration was practiced: after the cremation of the individual, the skeletal remains were placed into terracotta urns deposited inside rock cavity or in stone shelves provided with a lid (Di Salvo, 2004). In details, according to the German archaeologist Babette Bechtold, author of the opera called 'La necropoli di Lilybaeum' (1999), there are seven distinct funeral rituals

in Lilybaeum (extended on the present-day streets of Marsala city: 'via Cicerone', 'via Berta', 'via De Gasperi', 'Corso Gramsci', 'via Fante', 'via Cattaneo' and 'via F. Struppa'):

- **Ritual A**: inhumation ritual in pit grave cut in the rock (type I and II), discovered in 'via Cicerone', 'via Berta', 'via Cattaneo', and 'Corso Gramsci' city routes;
- **Ritual B**: *enchytrismòs* into pit grave 'type II', discovered in 'via A. De Gasperi';
- **Ritual C**: primary cremation ritual into pit grave 'type II', discovered in 'via Berta', 'via Cicerone' and A. De Gasperi streets;
- **Ritual D**: cremation ritual into urns (brought to light in 'Corso Gramsci' and 'via Cicerone' streets), or cremation on *ustrinum* (Di Salvo, 2004), near one or more pit graves (in 'via Cattaneo', 'via A. De Gasperi', 'Corso Gramsci', 'via Berta' and 'via F. Struppa' roads);
- **Ritual E**: primary cremation into pit grave 'type V' (into slabs), or 'type VI' (into pit earthy), often with *epitymbia* (in 'via Berta', 'via Cicerone' and 'via F. Struppa' streets);
- **Ritual F**: inhumation ritual into pit grave 'type V and VI' in 'via A. De Gasperi';
- **Ritual G**: secondary cremation into pit grave 'type V' an 'type VI' in 'via A. De Gasperi', 'via Berta' and 'via Del Fante' streets.

The material culture attested by grave goods reveals a deep Hellenization of the population living in Lilibeo, and the wideness of commercial networks in which the city was inserted, that surely beyond Greek Sicily, also had to reach the Southern Italy and the present-day Lazio area (Bondì, 2009).

The human remains analysed for this study were recovered in 'Corso Gramsci' street during the archaeological excavation of 2003-2004 (conducted by Dr Rossella Giglio and Dr Emanuele Canzonieri), carried out during the construction of the city sewage system in the North area of Marsala city. The archaeological campaign, conducted in according to the criteria of so-called 'preventive archaeology', has brought to light several tombs: n = 57 graves belonging to the common type of rectangular pit grave excavated in the rocky bench (some of them with the cremation ritual in situ or into cinerary urns); n = 1 *epitymbion* (pyramidal shape); n = 5 hypogeic tombs articulated in two types; one vertical shaft leading into two funerary chambers (S.U. 105 and S.U. 603), and one hypogeic chamber *tombs*

preceded by a *dromos* (corridor) (Giglio and Canzonieri, 2009). The sampling was carried out before the anthropological analysis conducted by Dr Jessica Sardo (Sardo, 2012), subject to authorization by the 'Soprintendenza BB. CC. e AA. di Trapani'.

5.2. Materials and Methods

5.2.1. Samples for genetic analysis

Sample	*Material*	*Site*	*Sample info*	*Years of excavation*	*Researchers* [a]
TH1	Long bone	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH2	Long bone	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH3	Tooth	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH5	Femur bone	Tharros	Grave A2, S.U. 188	2014	R1, R2, R3, R4
TH6	Femur bone	Tharros	Grave A2, S.U. 188	2014	R1, R2, R3, R4
TH7	Femur bone RS	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH8	Femur bone LS	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH9	Femur bone RS	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH10	Femur bone LS	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3, R4
TH11	Long bone	Tharros	Grave Z, S.U. 167	2014	R1, R2, R3
TH13	Tooth	Tharros	Grave A2, S.U. 179	2014	R1, R2, R3, R4
TH14	Tooth	Tharros	Grave Z, S.U. 167	2014	R1, R2, R3
TH15	Tooth	Tharros	Grave Z, S.U. 167	2014	R1, R2, R3
TH18	Femur bone RS	Tharros	Grave A2, S.U. 178	2014	R1, R2, R3
DA390	Tooth	Tharros	Grave B14, S.U. 72	2015	R1
DA391	Tooth	Tharros	Grave 2, S.U. 16	2012	R1, R2, R3
DA392	Petrous bone	Tharros	Grave L, S.U. 1001	2013	R1, R2, R3
DA393	Petrous bone	Tharros	Grave L, S.U. 1003	2013	R1, R2, R3
DA394	Petrous bone	Tharros	Grave L, S.U. 1003	2013	R1, R2, R3
DA395	Petrous bone	Tharros	Grave L, S.U. 1001	2013	R1, R2, R3
DA396	Tooth	Tharros	Grave 2, S.U. 16	2012	R1, R2, R3
DA397	Tooth	Tharros	Grave B9, S.U. 9003	2016	R1
DA398	Tooth	Tharros	Grave B10, S.U. 85	2015	R1
DA399	Tooth	Tharros	Grave B7, S.U. 34	2015	R1
DA400	Tooth	Tharros	Grave B9, S.U. 35	2015	R1
DA401	Tooth	Lilybaeum	Grave 159, S.U. 181	2003-2004	R1
DA402	Tooth	Lilybaeum	Grave 161, S.U.	2003-2004	R1
DA403	Tooth	Lilybaeum	Grave 718, Chamber 705	2003-2004	R1

Table 5.2.1.1 | Information about the Punic samples analysed. [a]Researchers who had been in contact with the ancient samples during the archaeological excavation and the laboratory work. Abbreviations: RS, right side; LS left side. For more detailes about grave A2 see Fariselli (2017). Archaeological data of graves B7, B9 and B14 are reported in Secci (2016).

A total of 28 samples were collected from two different localities in Sardinia (Tharros site) and Sicily (Lilibeo site) Islands (Table 5.2.1.1). Regarding the southern necropolis of Tharros,

the sampling was carried out throughout the archaeological campaigns ranging from 2012 to 2016, followed stringent *in-situ* procedures previously described in literature to reduce the risk of contamination and increase DNA preservation (Pruvost et al., 2007; Bollongino et al., 2008; Fortea et al., 2008). More details about contamination avoidance procedures are reported in "Sample for genetic analysis" section of Case Study I. The human specimens consisted of a large set of disarticulated and fragmented human bones, including long bones (femurs), fragments of the skull (petrous bones or maxillary bones) and loose teeth.

From Lilibeo site, samples were collected immediately after excavation field of 2003-2004. The specimens selected for this study were teeth of three different individuals, and the sampling was carried out in conditions designed to limit contamination with exogenous DNA, such as wearing protective equipment (surgical gown, face mask and gloves) and sterilizing the collecting material with DNA oxidant such as 5% sodium hypochlorite (NaClO).

5.2.2. Radiocarbon dating

Collagen is the dominant organic component of bone and is intimately locked within the hydroxyapatite structure of this ubiquitous biomaterial that dominates archaeological and palaeontological assemblages. Radiocarbon analysis of extracted collagen is one of the most common approaches to dating ancient bone samples (Harvey et al., 2016).

Selected bone samples (DA393 and DA394 from Tharros) were chemically prepared and measured by Accelerator Mass Spectrometry (AMS) by using a 3MV Tandetron accelerator at CEDAD-CEnter for DAting and Diagnostics (University of Salento), to determine the absolute chronology by using the radiocarbon 14C method. Punic samples were mechanically cleaned and grounded into fine powders, in order to remove surface contamination before chemical treatments. Then, the Longin method (Longin, 1971) modified by Brown et al. (1988) was used to extract collagen from the osteological samples. The bone powder (750 mg) was subjected by a sequence of washes with 0.5 M of HCl to decalcify the bone, 0.1M of NaOH (30 minutes) to remove humic substance and followed by 0.5M of HCl (15 minutes) at room temperature. Interspersed with rinsing with Ultrapure MilliQ™ water between each reagent. Then, 'crude collagen' was gelatinised in pH 3 solution at 75° C for 20 hours and filtered by using 0.45 mm pore silver filter. The purified collagen gelatine was then combusted to CO_2 in a sealed quartz tube, and reduced to graphite using iron powders as a catalyst in an excess H_2 atmosphere at 600° C. Finally, the graphite obtained from the reduction process was pressed into Al target and measured by AMS.

The radiocarbon dates of the samples were determined by measuring the 12C, 13C, 14C isotopes, corrected for the fractionation effects and normalized by using the IAEA C6 Sucrose standards. The OxCal v.3.10 programme (Bronk Ramsey 1995, 2001) and the INTCAL09 calibration curve (Reimer et al., 2013) were used to calibrate the radiocarbon data into calendar ages at 1s and 2s confidence level.

5.2.3. Ancient DNA procedures

Genetic analyses of 14 Punic samples were performed at the Laboratories of Physical Anthropology and Ancient DNA, Department of Cultural Heritage (DBC), University of Bologna. aDNA laboratory standards employed to avoid contaminations (Cooper and Poinar, 2000; Fulton et al., 2012; Knapp et al., 2012, 2015) were described in "Ancient DNA procedures" section of Case Study I. PCR run and post-PCR laboratory procedures were carried out in a separate building at the Laboratory of Molecular Anthropology and at the Centre for Genome Biology, Department of Biological, Geological and Environmental Sciences (BiGeA), University of Bologna. DNA extraction and sequencing library construction steps of the remaining ancient samples (n=14) were performed at the Centre for GeoGenetics, Copenhagen, Denmark, in ancient DNA facilities dedicated to the analysis of ancient hominin samples and physically separated from post-PCR and modern DNA laboratories (Willerslev and Cooper, 2005; Gilbert et al., 2005). As a result, the samples were not all processed in the same manner, but by two different methodological approaches outlined below.

5.2.4. Molecular analysis I

5.2.4.1. Cleaning and powdering

Bone and tooth samples (Figure 5.2.4.1.1) were decontaminated by removing the outer surface through sterile blades or by a diamond pointy drill-bit with a Dremel® drill (Dremel, Racine, WI, USA). Additionally, the surface of the teeth was gently wiped with 5% sodium hypochlorite (NaClO) and rinsed with nuclease-free water. Bone and tooth samples were then UV-irradiated (254 nm wavelength, 12 V and a distance of 5 cm from the UV source) in a cross-linker for 60 minutes from each side, subsequently ground to a fine powder with a mortar and stored at 5 °C until use. All metallic material was thoroughly cleaned with a

bleach solution after use, rinsed with 70% ethanol and UV irradiated for 15 minutes to avoid cross-contamination between samples (for more detail see “Sample preparation” section of Case Study I).

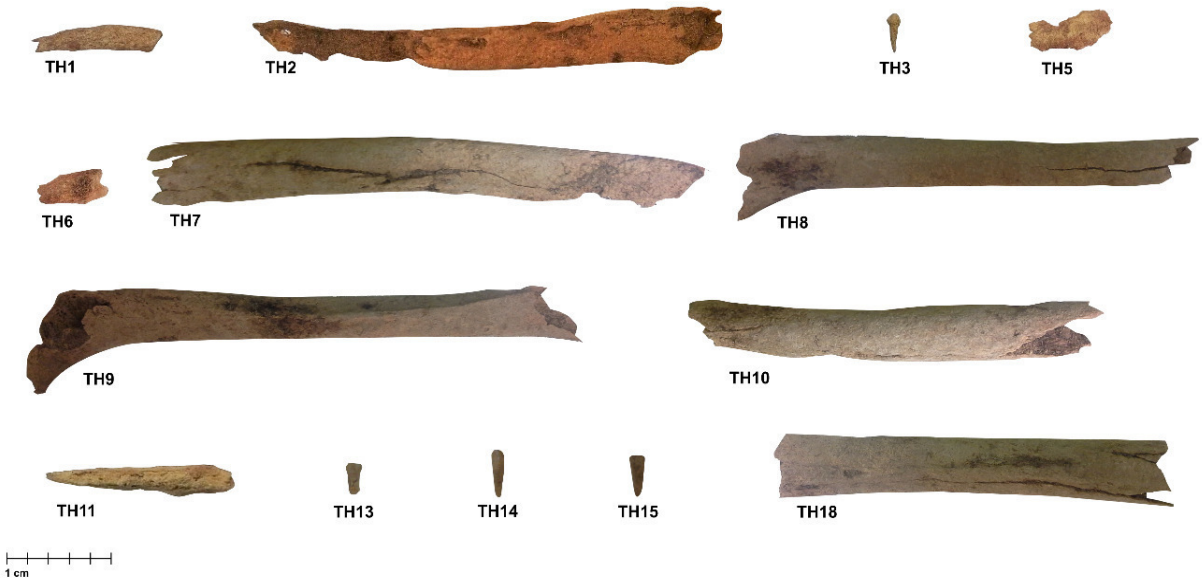


Figure 5.2.4.1.1 | Osteological remains from Tharros necropolis analysed in the present study (molecular analysis I).

5.2.4.2. Ancient DNA isolation

Before DNA extraction, the drilled bone material (ranging from 100 to 300 mg) was digested in 1 mL proteinase K (0.25 mg/mL) and EDTA (0.45M EDTA) lysis buffer and DNA was extracted through a silica-based method (Dabney et al., 2013) slightly modified (see “Ancient DNA extraction” section of Case Study I). Briefly, after centrifugation for 2 min at 13,000 rpm, the supernatant mixed with 10 mL of binding buffer (5M guanidine hydrochloride, 40% isopropanol, 0,05% Tween-20, 90 mM sodium acetate and nuclease-free water) was transferred on a Zymo-Spin™ V reservoir (Zymo Research -Irvine, CA, USA) previously treated with bleach and UV light to avoid contamination, fitted on a MinElute column (Qiagen GmbH, Hilden, Germany). After a centrifugation step, the MinElute column was then placed in a collection tube and centrifuged at 6,000 g for 1 min. The column was washed twice with 750 µL of PE buffer (Qiagen GmbH, Hilden, Germany), and the DNA was eluted in 35 µL TET buffer (1 mM EDTA pH 8, 10 mM Tris-HCl pH 8, 0.05% Tween-20). All extracts were stored at -20°C in siliconized tubes. The DNA concentration in all final

extracts was measured using Qubit® Fluorometric Quantitation (Life Technologies, Grand Island, NY).

5.2.4.3. Real-Time PCR

The Real-Time PCR allows amplifying the DNA template but at the same time, it monitors the developing of the reaction, in order to make a relative quantification. This is possible thanks to fluorescent markers, that follow the same reaction kinetics of the PCR. The emitted fluorescence (caused by a specific irradiation coming from thermocycler) is measured in real time from a CCD camera and all the measurement operations are managed by a specific software.

The quantification of the collected samples was carried out with the Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Oyster Pint, CA) according to manufacturer's early kit evaluation instruction. The Quantifiler® Trio Kit detects three different human multi-copy autosomal targets producing a small amplicon (SA - small autosomal target), a larger amplicon (LA - large autosomal target) and Y-Chromosome target. The SA target is the primary quantification target for total human genomic DNA. Its smaller amplicon size (80 bp) is aligned with the sizes of typical "mini" STR loci and makes it better able to detect degraded DNA samples. The LA target is used mainly as an indicator of DNA degradation (DI), by comparing the ratio of its quantification result with that of the SA target. The Y target allows the quantification of a sample's human male genomic DNA component and is particularly useful in assessing mixture samples of male and female genomic DNAs. Moreover, an internal PCR control (IPC) allows checking the presence of inhibitors in the reaction mix.

The standard curve was generated with tenfold serial dilutions obtaining 5 standard concentrations from 50 ng/µL to 0.005 ng/µL, and duplicate reactions of each standard curve sample were run per plate. Each Real-Time PCR amplification reaction contained 8 µL of Quantifiler® Trio Primer Mix and 10 µL of Quantifiler® Trio PCR Reaction Mix and 2 µL of sample DNA. Quantitative PCR was performed using the 7500 Real-Time PCR System (Thermo Fisher Scientific, Oyster Point, CA) with 96-well Optical MicroAMP plates following the manufacturer's instruction. The data were analysed using the HID Real-Time PCR Analysis Software v1.2 using the default setting provided with the Quantifiler® templates.

5.2.4.4. Analysis of mtDNA control region

Three sets of PCR primers pairs (L15995-H16132: 179 bp, L16107-H16261: 197 bp, L16247-H16402: 156 bp) (Caramelli et al., 2003) were used to amplify overlapping DNA fragments of the first hypervariable segment (HVS-I) of the mtDNA control region, to obtain 360 bp, spanning from nucleotide position (np) 16024 to np 16383. The amplification of each fragment was carried out in independent PCR reactions. For each ancient sample at least two extractions were undertaken at different time points, two amplifications for each extraction were made and both strands of the DNA were sequenced, to assess the reproducibility of the results (Ottoni et al., 2011; Hervella et al., 2015; Lorkiewicz et al., 2015). Moreover, all DNA extracts were screened to test their appropriate molecular behaviour (Cooper and Poinar, 2000) with L15996-H16401 primers pairs (~400bp) (Vigilant et al., 1989), in order to detect possible contaminations, given that ancient DNA molecules are often fragmented into very short pieces (60-150 bp) (Prüfer et al., 2010). Details about the amplification reaction are reported in "mtDNA amplification" section of Case Study I.

5.2.4.5. Typing of mtDNA coding region SNPs

Twenty-two haplogroup-diagnostic mtDNA coding region SNPs (4216L, 4529L, 4580L, 7028L, 10398L, 10400L, 10873L, 12308L, 12705L, 14766L, 3010L, 3915H, 3936H, 3992L, 4310L, 4745L, 4336L, 4769H, 4793H, 6776H, 13708L, 13759L) (Richards et al., 2000; Herrnstadt et al., 2002) were typed to confirm the haplogroup assignment preliminarily inferred with the HVS-I haplotype motifs. The genotyping has been performed by means of two different multiplex-PCR reactions followed by a single-base extension assay carried out with the SnaPshot® Multiplex Kit (Applied BioSystems, Foster City, USA) (Bertoncini et al., 2011). Capillary electrophoresis reaction was performed at the Department of Diagnostic and Laboratory Services and Legal Medicine (University of Modena and Reggio Emilia) on an ABI PRISM™ 3130 DNA Genetic Analyzer (Applied BioSystems, Foster City, USA). Details about the amplification and multiplex PCR reactions are reported in "SNPs genotyping" section of Case Study I.

5.2.4.6. Sanger sequencing

Sanger sequencing experiment was performed on a 3730 DNA Analyzer (Applied BioSystems, Foster City, USA) at the "Unità Operativa di Genetica Medica dell'Azienda Ospedaliera di Bologna", with collaborating researchers.

To perform the sequencing reaction between 0.5-1.5 μL of purified products were used along with the Byg Dye Terminator Mix. This last is constituted by Taq polymerase and dideoxynucleotides (ddNTPs) marked with four different fluorochromes (one for each different base). See "Sanger sequencing" section of Case Study I for details about the procedure and the composition of the reaction mix. Briefly, a sample plate was prepared, uploading each well with reaction mix and purified DNA. This plate was then positioned in the Thermal Cycler.

After the modified amplification reaction, the products were purified with a mix made of nuclease-free water, ethanol 100%, sodium acetate, and several steps of centrifugation. This precipitation reaction is aimed to the elimination of ddNTPs not incorporated during the amplification reaction: the ethanol links with DNA (thanks also to centrifugation) and brings it at the bottom of the well plates, while the supernatant is discarded. This procedure is repeated two times to obtain the optimal pureness of samples. The output of automatic sequencing is an electropherogram represented by a series of peaks, each one corresponding to a distinct nucleotide and indicated with a featuring colour and letter (A green, T red, C blue and G black). This chromatogram was converted in a bases-sequence through the Chromas 2.6.4 software (Technelysium).

5.2.4.7. DNA analysis of the researchers

Mitochondrial DNA genotypes of the researchers who handled the ancient specimens were determined. The epithelial cells were collected from the mucosa on both sides of the oral cavity using buccal swabs. DNA extraction, as well as PCR and sequencing reaction setup involving modern samples, was carried out in at the laboratory of Molecular Anthropology (University of Bologna) that was physically separated from the laboratory where the ancient samples were analysed. DNA was extracted using QIAmp DNA Mini Kit (Qiagen GmbH, Hilden, Germany) enzymatically lyses the cells and, again, uses spin columns with a silica gel membrane that selectively binds nucleic acids and allows their purification from cellular residuals. The entire mtDNA HVR-I region was amplified using a single pair of primer (L15996-H16401) reported in Vigilant et al. (1989). The amplification products were purified with the MinElute PCR purification Kit (Qiagen GmbH, Hilden, Germany), and then sequenced directly with the same amplification primers (forward and reverse) following the BigDye Terminator v1.1 Cycle Sequencing Kit supplier's instructions. Sanger sequencing

experiment was performed on a 3730 DNA Analyzer instrument (Applied BioSystems, Foster City, USA).

5.2.4.8. Statistical analyses

5.2.4.8.1. Intra-population analysis

5.2.4.8.1.1. *Haplogroup assignment*

The obtained data - three HVS-I sequences (FWR and RVS) for each extracted sample - were aligned and compared to the revised Cambridge Reference Sequence (rCRS, GenBank Accession Number NC 012920) (Anderson et al., 1981; Andrews et al., 1999) using BioEdit v7.2.5 (Hall, 1999) and MEGA7 software (Kumar et al., 2016), to define the HVR-I mutational motif. The types and frequency of nucleotide variations among the sequences were checked, such as C->T transitions, which represent the prevalent signal of post-mortem miscoding lesions in authentic aDNA (Stiller et al., 2009; Bollongino et al., 2013; Dabney at al., 2013b). Haplogroups assignment was determined using Haplogrep2 software (Kloss-Brandstätter et al., 2011) with a further check based on the PhyloTree mtDNA phylogeny, built 17 (www.phylotree.org) (van Oven and Kayser, 2009).

5.2.4.8.1.2. *Summary and population differentiation statistics*

To evaluate the intra-population genetic variability, a summary statistics analysis was computed with the Arlequin software ver. 3.5 (Berne, Switzerland) (Excoffier and Lischer, 2010):

- Number of different haplotypes (k)
- Gene diversity (H)
- Mean number of pairwise differences (π)
- Nucleotide diversity (π_n)

Information regarding the parameters analysed is provided in 'Summary and population differentiation statistics' section of Case Study I.

5.2.4.8.2. Inter-population variation

5.2.4.8.2.1. *Additional populations used in comparative analyses*

To perform comparative and phylogenetic analyses, 3258 mtDNA sequences (nps 16024-16383) belonging to 41 present-day Mediterranean populations were collected from the

literature (Table 5.2.4.8.2.1): Europe (n = 1335 - Cyprus, France, Greece, Italy, Spain, and Portugal); North Africa (n = 1088 - Algeria, Libya, Morocco, and Tunisia); Near East (n = 835 - Jordan, Lebanon, Palestine, and Syria).

Populations	***Sample size***	***Code***	***References***
EUROPE			
ITALY			
Cabras, Sardinia	48	CB	*Sanna et al. 2011*
Medio Campidano, Sardinia	23	CM	*Sanna et al. 2011*
Oristano, Sardinia	39	OR	*Boattini et al. 2013*
Olbia/Nuoro, Sardinia	31	OTNU	*Boattini et al. 2013*
San Pietro, Sardinia	44	SP	*Falchi et al. 2006*
Sant'Antioco, Sardinia	42	SA	*Falchi et al. 2006*
Nuoro, Sardinia	51	NUO	*Falchi et al. 2006*
Gallura, Sardinia	50	GAL	*Falchi et al. 2006*
Trexenta, Sardinia	47	TX	*Falchi et al. 2006*
Ragusa/Siracusa, Sicily	39	RGSR	*Boattini et al. 2013*
Catania, Sicily	37	CT	*Boattini et al. 2013*
Agrigento, Sicily	42	AG	*Boattini et al. 2013*
Trapani, Sicily	40	TP	*Brisighelli et al. 2012*
SPAIN			
Granada	121	GRA	*Hernández et al. 2014*
Huelva	158	HUE	*Hernández et al. 2014*
Majorca	67	MAJ	*Falchi et al. 2006*
PORTUGAL			
Northern	100	POR_N	*Pereira – Prata - Amorim 2000*
Central	82	POR_C	*Pereira – Prata - Amorim 2000*
Southern	59	POR_S	*Pereira – Prata - Amorim 2000*
FRANCE			
Corsica	53	COR	*Falchi et al. 2006*
GREECE			
Greek from Attica	29	GR_A	*Varnesi et al., 2001*
Greek from Thrace	25	GR_T	*Bosch et al., 2006*
Greek from Crete	18	GR_C	*Varnesi et al., 2001*
CYPRUS			
Cypriots	90	CYP	*Irwin et al. 2008*
Total European sequences	***1335***		
NORTH-AFRICA			
LIBYA			
Libyans	269	LYB	*Fadhlaoui-Zid et al. 2011*
Fezzan (Al Awaynat)	111	ALA	*Ottoni et al. 2009*
Fezzan (Tahala)	18	TAH	*Ottoni et al. 2009*

Table 5.2.4.8.2.1 | Data collected from 41 current Mediterranean populations (continued).

Populations	*Sample size*	*Code*	*References*
MOROCCO			
Moroccans	149	MOR	*Behar et al. 2008*
TUNISIA			
Qalaat El Andalous	29	QEA	*Cherni et al. 2009*
Capital Tunis	51	TS	*Cherni et al. 2009*
El Alia	48	AL	*Cherni et al. 2009*
Zriba	34	ZRI	*Cherni et al. 2009*
Slouguia	27	SLA	*Cherni et al. 2009*
Testour	49	TES	*Cherni et al. 2009*
Skira	20	SK	*Cherni et al. 2009*
Kesra	43	KES	*Cherni et al. 2009*
ALGERIA			
Algerians	240	ALG	*Bekada et al. 2013*
Total North-Africa sequences	***1088***		
NEAR EAST			
PALESTINE			
Palestinian Israel	211	PAL	*Richards et al. 2000; Behar et al. 2008*
SYRIA			
Syrians	106	SYR	*Behar et al. 2008*
JORDAN			
Jordan	145	JOR	*González et al. 2008*
LIBANON			
Lebanese	363	LEB	*Heber et al. 2011*
Total Near-East sequences	***835***		

Table 5.2.4.8.2.1 | Data collected from 41 current Mediterranean populations.

5.2.4.8.2.2. *Genetic Distances*

Population-specific pairwise genetic distances (F_{ST}) were calculated with Arlequin software version 3.5.1.2 (Berne, Switzerland) (Excoffier and Lischer, 2010) using a Kimura's two-parameter model (Kimura, 1980) and a gamma distribution of 0.26.

5.2.4.8.2.3. *Multidimensional scaling (MDS)*

Slatkin Fst-values (Slatkin, 1995) between Punic samples and comparison dataset samples were used to reconstruct non-metric Multidimensional scaling plot visualized in bi-dimensional space using R 'MASS' package (R-DevelopmentCoreTeam 2008).

5.2.4.8.2.4. *Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set (Jolliffe, 2002).

It accomplishes this reduction by identifying directions, called principal components (PCs), along which the variation in the data is maximal. By using a few components, each sample can be represented by relatively few numbers instead of by values for thousands of variables. Samples can then be plotted, making it possible to visually assess similarities and differences between samples and determine whether samples can be grouped (Ringnér, 2008).

PCA analysis was performed using the R *adegenet* package (Jombart, 2008).

5.2.5. Molecular analysis II

5.2.5.1 Sample preparation

Figure 5.2.5.1.1 | Osteological remains from Tharros and Lilybaeum necropoleis analysed in the present study (molecular analysis II). Black line: 1 cm.

To minimize potential contamination from modern DNA due to the previous handling, the surface of the samples (Figure 5.2.5.1.1) was first removed inside a UV cabinet. All teeth samples were cleaned mechanically by gently running a sterile disc drill over the outermost surface. Then, the cementum in the outer layer of the root was target by removing as much dentine as possible according to Damgaard et al. (2015). In poorly preserved samples, where the cementum was absent or partly absent, the outer layer of the roots was taken as recommended in Hansen et al. (2017). The petrous bones were sampled by cutting off the

apex part, slice by slice, until the otic capsule was reached (Pinhasi et al., 2015, Hansen et al., 2017), since the denser part gave higher endogenous DNA yields (Gamba et al., 2014, Pinhasi et al., 2015). All the samples were then cut into small pieces (not pulverised) and stored in a 5-mL tube at 4 °C until use. All tools used in this phase were consecutively cleaned with the following products: 5% NaCl, DNA AWAY™ Surface Decontaminant (ThermoFisher), 96% ethanol and then exposed to UV-irradiated for 30 minutes on each side. The amount of starting material varied but was generally 300-700 mg.

5.2.5.2 DNA extraction

All the samples were 'pre-digested' in 5 mL of extraction buffer (Table5.2.5.2.1) for 15-30 minutes at 37°C under constant rotation. After centrifugation for 2 minutes at 16,000 g, the supernatant was discarded. An identical fresh extraction buffer was transferred again to the pre-digested and sedimented pellet, then the sample was vortexed and incubated for a full overnight at 37°C. This treatment facilitates the removal of surface contaminants and hence results in a higher proportion of endogenous DNA in the final extract (see Damgaard et al., 2015 and Allentoft et al., 2015).

Reagent	***mL/µL per sample***
0.5M EDTA solution	4.7 mL
0.14-0.22 mg/mL Proteinase K solution	50 µL
10% N-Laurylsarcosyl (or Tween 20)	250 µL
Total	5 mL

Table5.2.5.2.1 | Extraction buffer.

Following incubation, samples were centrifuged for 2 minutes at 13,000 rpm and the supernatant was transferred to a new 50-mL tube. The DNA was extracted using a silica-powder-based extraction method, by means of a binding buffer described in Allentoft et al. (2015), which is tailored to ultra-short DNA fragments. The silica suspension (Table5.2.5.2.2) was prepared by mixing SiO_2 with nuclease-free water, followed by 1 hour of sedimentation. Afterwards, 48 mL of supernatant was transferred to a 50-mL tube followed by another 5-hour of sedimentation. Then the top 43 mL was carefully removed, and the silica was resuspended and activated with 60 µL 37% HCl for use.

Material	*Amount per 5 mL*
SiO_2	6 g
Nuclease-free water	50 mL

Table 5.2.5.2.2 | Silica suspension.

For the DNA isolation, 45 mL of binding buffer (Table 5.2.5.2.3) and 100 µL of silica suspension were transferred to each sample and adjusted to pH 4-5 with 37% HCl, followed by incubation on a rotor (~ 30 rpm) at room temperature for 1-hour in the dark.

Reagent	*Amount per ~11 samples*
Qiagen buffer PB	500 mL
5M Sodium Acetate	9 mL
5M Sodium Chloride	2.5 mL
37% HCl	2 mL
Qiagen pH indicator	500 µL

Table 5.2.5.2.3 | Binding buffer.

After that, the sample was centrifuged for 2 minutes at 5000 g and the supernatant was removed. The pelleted silica was resuspended by pipetting up and down and transferred into a fresh 2-mL tube, to makes handling easier for the following steps. The tube was centrifuged for 15 seconds at 16,000 g, the supernatant was discarded, and the silica-pellet was washed twice with 680 µL of 80% cold ethanol (store at -20°C). The silica-pellet was centrifuged again for 15 seconds at 16,000 g, the supernatant was carefully discarded, and the silica-pellet was dried at room temperature for ~1 minute with open lids. The dried silica was then resuspended in 82 µL of EB buffer (Qiagen GmbH, Hilden, Germany) by stirring through a pipette tip. Finally, the sample was centrifuged for 2 minutes at 16,000 g, and 80 µL supernatant was stored at -20 °C into a new 1,5-mL siliconized tube until use.

Each extraction session included a mock extraction blank where no bone or tooth powder is added to the extraction reagents.

5.2.5.3 Library preparation

Following extraction, DNA template was built into a blunt-end library using the NEBNext DNA Library Prep Master Mix Set for 454 (NEB reference E6070) (New England Biolabs Inc) with 10 pM Illumina multiplex adapter (Meyer and Kircher, 2010), according to manufacturer's instructions, with a few modifications described below. The initial

nebulization step was omitted since aDNA is already fragmented. Negative control was included in each PCR batch.

5.2.5.3.1 *Blunt-end repair*

A volume of 20 µL of DNA extracted was incubated with a blunt-end repair reaction (Table 5.2.5.3.1.1) containing NEBNext End Repair Reaction Buffer (10X) and NEBNext End Repair Enzyme Mix (New England Biolabs, Inc.), which removes the 3' end overhangs, synthesizes complementary strands to the 5'end overhangs, leaving the DNA blunt-ended with 5' end phosphates and 3' end hydroxyls. The reaction was incubated for 20 minutes at 12°C and 15 minutes at 37°C. After cycling, blunt-end DNA was purified using the MinElute silica spin-columns (Qiagen GmbH, Hilden, Germany), with 250 µL of 10X volume manufactured PB buffer (Qiagen GmbH, Hilden, Germany) and 650 µL of PE Buffer (Qiagen GmbH, Hilden, Germany) according to manufacturer's instructions. DNA was then eluted in 16 µL of EB Buffer (Qiagen GmbH, Hilden, Germany).

Reagent	***Volume per 1 sample (µL)***
NEBNext 10x end repair buffer	2,5
NEBNext end repair enzyme mix	1,25
Fragmented DNA sample	20
Total	23,75

Table 5.2.5.3.1.1 | Blunt-end repair.

5.2.5.3.2 *Adapter ligation reaction*

Illumina-specific adapters were prepared as in Meyer and Kircher (2010) (Table 5.2.5.3.2.1) and ligated to the end-repaired DNA using T4 DNA ligase in 25 µl reactions mix (Table 5.2.5.3.2.2). The reaction was incubated at 20°C for 15 minutes. DNA was then purified using the MinElute silica spin-columns (Qiagen GmbH, Hilden, Germany), with 125 µL of 5X volume manufactured PB buffer (Qiagen GmbH, Hilden, Germany) and 650 µL of PE Buffer (Qiagen GmbH, Hilden, Germany). Finally, DNA was eluted in 21 µL of EB Buffer (Qiagen GmbH, Hilden, Germany).

Illumina adapters ID	*Sequence*
IS1_adapter_P5	5'-AATGATACGGCGACCACCGA
IS2_adapter_P7	5'-CAAGCAGAAGACGGCATACGA
IS3_adapter_P5+P7	3' AGATCGGAAGAGC

Table 5.2.5.3.2.1 | Illumina adapters.

Reagent	*Volume per 1 sample (µL)*
Quick ligation 5x buffer	5
Illumina DNA adaptors (25µM stock)	0,5
Quick T4 Ligase	2,5
Nuclease-free water	2
End-repaired DNA sample	15
Total	25

Table 5.2.5.3.2.2 | Adapter ligation.

5.2.5.3.3 *Bst-DNA polymerase fill-in reaction*

In a volume of 25 µL, Illumina adaptors were ligated to the DNA 5' end, and a Bst Polymerase (New England Biosciences) was used to fill-in the sequence between the 3'end adaptors and the inserted DNA. The reaction (Table 5.2.5.3.3.1) was incubated at 37°C for 20 minutes followed by an increase to 80° C for 20 minutes to denature the Bst Polymerase.

Reagent	*Volume per 1 sample (µL)*
Adapter fill-in reaction buffer	2.5
Bst DNA polymerase	1.5
dNTPs (10 mM each)	1
Adapter-ligated DNA sample	20
Total	25

Table 5.2.5.3.3.1 | Fill-in reaction.

5.2.5.4 Library PCR amplification

The DNA library was then amplified and indexed in a 50 µL PCR reaction containing 1X KAPA HiFi HotStart Uracil+ReadyMix (KAPA Biosystems, Woburn, MA, USA) and 200 nM of each of Illumina's Multiplexing PCR inPE1.0 forward primer (5'AATGATACGG CGACCACCGA GATCTACACT CTTTCCCTAC ACGACGCTCT TCCGATCT) and a custom designed index reverse primer (5'CAAGCAGAAG ACGGCATACG AGATNNNNNN GTGACTGGAG TTC, where N's correspond to a 6-nucleotide index tag) (Tables 5.2.5.4.1a, 5.2.5.4.1b). The DNA

amplification and all subsequent steps were performed in a standard molecular laboratory at the University of Copenhagen.

Reagent	*Volume per 1 sample (µL)*
1X KAPA HiFi HotStart Uracil+ ReadyMix	25
Illumina Multiplexing PCR primer inPE 1.0	1
Illumina Index PCR primer	1
3' end modified DNA sample	11
Nuclease-free water	12
Total	50

Table 5.2.5.4.1a | Library amplification reaction.

Sample ID	*Code index*	*Index*
DA390	13	CTATCA
DA391	25	TAGATG
DA392	17	CGCTAT
DA393	18	TGAACA
DA394	20	CAGCTA
DA395	16	GTGTAT
DA396	33	GACCGG
DA397	21	ACATAC
DA398	41	TGTCTG
DA399	24	CGATGA
DA400	32	ACGCAT
DA401	19	GTATCT
DA402	28	TCTCGC
DA403	22	TGAGCC

Table 5.2.5.4.1b | Nucleotide sequence of indexes used in this study.

The first amplification was carried out with initial 1 minutes at 98°C, followed by 12 cycles of 15 seconds at 98°C, 30 seconds at 65°C and 30 seconds at 72°C, ultimately with a 1 minute of elongation step at 72°C.

Libraries were subjected to a second PCR amplification round using 5 µL of the 'pre-amplified' library and P5 and P7 primers (Meyer and Kircher, 2010) (Table 5.2.5.4.2) and using 8 cycles in the amplification step.

Reagent	*Volume for one reaction (µL)*
1X KAPA HiFi HotStart Uracil+ ReadyMix	12.5
P5 primer	1
P7 primer	1
Pre-amplified sample	5
Nuclease-free water	5
Total	25

Table 5.2.5.4.2 | Library re-amplification reaction.

The amplified library was purified using the MinElute silica spin-columns (Qiagen GmbH, Hilden, Germany), with 125 µL of 5X volume manufactured PB buffer (Qiagen GmbH, Hilden, Germany) and 650 µL of PE Buffer (Qiagen GmbH, Hilden, Germany). Finally, DNA was eluted in 21 µL of EB Buffer (Qiagen GmbH, Hilden, Germany). Then, 2 µL of purified product was visualised following an electrophoresis on a 3,5% agarose gel and finally library concentration and size were checked on a 2100 Bioanalyzer using the High Sensitivity Kit (Agilent Technologies, Palo Alto, CA, USA). None of the extraction blanks or PCR blanks showed the presence of DNA and were therefore not further sequenced.

5.2.5.5 High-throughput sequencing and data analyses

Libraries were pooled at equimolar concentrations and were 'shot-gun' sequenced on an Illumina HiSeq™ 2500 platforms at the Danish National High-Throughput DNA Sequencing Centre (University of Copenhagen, Denmark) using 100 bp single read chemistry.

The Illumina data was base-called by means of Illumina software CASAVA 1.8.2 and sequences were de-multiplexed with a requirement of full match of the 6-nucleotide index that was used for library preparation (Table 5.2.5.4.1b). Quality features of the sequencing data were evaluated at different stages using FASTQC version 0.11.5 (http://www.bioinformatics.babraham.ac.uk).

Reads were trimmed, and adapters were removed using AdapterRemoval version 1.5.4 (Lindgreen, 2012), with a minimum trim length of 30 bp and a minimum base quality of 20. Merged reads were mapped against the human reference genome Build 37, available from the UCSC genome browser, using Burrows-Wheeler Aligner (BWA) version 0.6.2 (Li and Durbin, 2009) with the *samse* function using standard parameters except that seeding was disabled, to allow for higher sensitivity as recommended by Schubert et al. (2012).

Mapped reads were filtered for mapping quality 30 and sorted using the MergeSamFiles Picard tool (http://picard.sourceforge.net) and Samtools (Li, et al., 2009). Data was merged to

library level and duplicates removed using Picard MarkDuplicates (http://picard.sourceforge.net) and hereafter merged to sample level. Sample level BAMs were then realigned using GATK version 2.2.3 (DePristo et al., 2011). The relevant summary statistics were extracted with a custom Perl script.

5.2.5.6 Estimation of the contamination and authentication of data

DNA degrades over time and aDNA can, therefore, be characterised by certain types of damages that are not expected to be present in modern DNA. Abasic sites, strand breaks, interstrand cross-links and a wide diversity of atypic nucleotidic bases are formed following oxidative and hydrolytic degradation (Lindahl, 1993; Pääbo et al., 2004), even in the most favourable preservation conditions. Using standard parameters in the Bayesian approach implemented in mapDamage 2.0 (Ginolhac et al. 2011; Jònsson et al., 2013) typical characteristic damage patterns of the aDNA were assessed: *(i)* the frequency of C → T transitions at the first position at the 5' end of reads*, (ii)* λ, the fraction of bases positioned in single-stranded overhangs, *(iii)* δs, the estimated C → T transition rate in the single-stranded overhangs. Outputs from mapDamage 2.0 were analysed and plotted with R.

Although the continuous implementation of the aDNA protocols, it is difficult to completely avoid contamination from modern DNA when working with ancient human material. Given this, it is necessary to establish that the endogenous DNA content obtained is not an effect of modern DNA contamination. For samples having mitochondrial genomes with >5X coverage, we determined contamination levels using the contamMix software (Fu et al., 2013) by comparing mapping affinities of each read to the consensus mitogenome of the ancient sample with mapping affinities to 311 mitogenomes worldwide as was done in Fu et al. (2013). Consensus sequences were determined using ANGSD (Korneliussen et al., 2014) using only sites with a minimum depth of 5X, and the 7 bp at the extremities of the reads were disregarded for the contamMix analyses to reduce the biases introduced by DNA damage.

5.2.5.7 Population genetics and statistical analyses

5.2.5.7.1 Population genetics analysis datasets

Datasets for population genetics analyses were constructed by merging ancient DNA data generated in this study as well as previous studies (Mathieson et al. 2015, Fu et al. 2016, Lazaridis et al. 2016, Lazaridis et al. 2017) with a reference panel of modern populations extracted from the literature. Genotypes for Punic ancient individuals were obtained at all

variant positions in the reference panel, discarding variants where alleles for the ancient individuals did not match either of the alleles observed in the panel. Since the Punic data obtained as well as the majority of the previously published ones are low coverage data, the 'mpileup' command of samtools (Li et al., 2009) (https://github.com/samtools/samtools) were used to extract reads overlapping the variants, then randomly sampling a single read with mapping quality ≥ 30 and base quality ≥ 30.

The modern reference dataset consists of 1753 contemporary individuals from 82 populations (S-Table 5.2.5.7.1.1a) genotyped for a common set of 279,801 autosomal SNPs. To avoid strand-flipping issues, ambiguous A/T and C/G polymorphisms were removed during the merging procedure. The obtained dataset was filtered using the PLINK software 1.07 (Purcell et al. 2007) to include only SNPs with genotyping success rate higher than 98% and minor allele frequency higher than 1%, and by removing individuals showing more than 1% of missing genotypes. In addition, cases of genetic relatedness among samples were tested by estimating the degree of identity-by-descent (IBD) sharing and by excluding one individual for each pair of samples with kinship coefficient (PiHat) higher than 12.5% (3rd degree relatives). The filtered modern dataset was finally merged with the Punic ancient sample as well as with available literature data for 390 ancient individuals (S-Table 5.2.5.7.1.1b). After the merging, we obtained a set of 267,391 SNPs genotyped in 2144 individuals.

We thinned the dataset with the PLINK software, by excluding SNPs in strong LD ($r2 > 0.1$) within a sliding window of 50 SNPs advanced by 10 SNPs at the time.

5.2.5.7.2 Principal component analysis (PCA)

As a first assessment of the genetic affinities of the Punic samples, a principal component analysis (PCA) was carried out using the *smartpca* function implemented in the EIGENSOFT package v6.0.1 (Patterson et al., 2006). The dataset was thinned with the PLINK software, by excluding SNPs in LD ($r2 > 0.4$) within a sliding window of 200 SNPs advanced by 25 SNPs at the time (--indep-pairwise 200 25 0.4). The PCA analysis was performed on the set of reference modern populations, and ancient individuals were subsequently projected onto the inferred PCA space by means of the lsqproject = YES function.

PCA results were plotted using a custom R script.

5.2.5.7.3 Outgroup -f_3 statistics

Outgroup-f3 statistic in the form of f_3 (*Mbuti; Tharros_Punic, X*) was used to formally assess the relationships between the Punic sample and the other ancient population groups. This statistic (Raghavan et al., 2013) is indeed expected to be proportional to the amount of shared genetic drift between $Population_1$ and $Population_2$ from a common ancestor. Unlike methods based on pairwise genetic distances such as F_{ST}, the F-statistics is less affected by genetic drift specific of $Population_1$ or $Population_2$. The outgroup f3-statistic was computed with the *qp3pop* function of the ADMIXTOOLS package (Patterson et al. 2012), using the Mbuti Pygmies as outgroup.

5.3 Results and Discussion

5.3.1 Radiocarbon dating

The degree of organic preservation remains a crucial concern when radiocarbon dating is applied to any bone material (Hedges et al., 1992), and the preservation state depends on two main aspects: *(i)* time, and *(ii)* physical/biotic environment; or more likely, a combination of the two (Harvey et al., 2016).

Radiocarbon dates were successfully obtained from the ancient bone material coming from the necropolis of Tharros, DA393 and DA394 samples, revealing ages ranging from 382 ± 153 BC to 364 ± 90 BC respectively (Figure 5.3.1.1a, 5.3.1.1b).

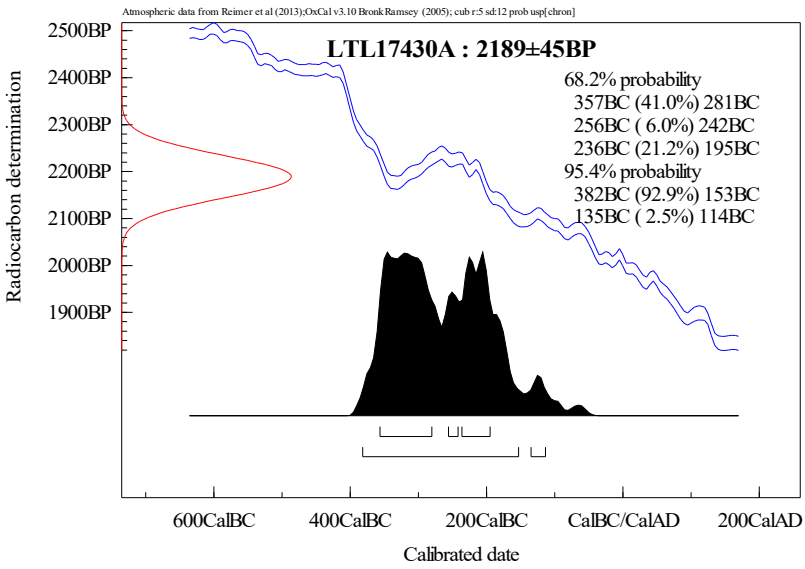


Figure 5.3.1.1a | Calibrated radiocarbon age of DA393.

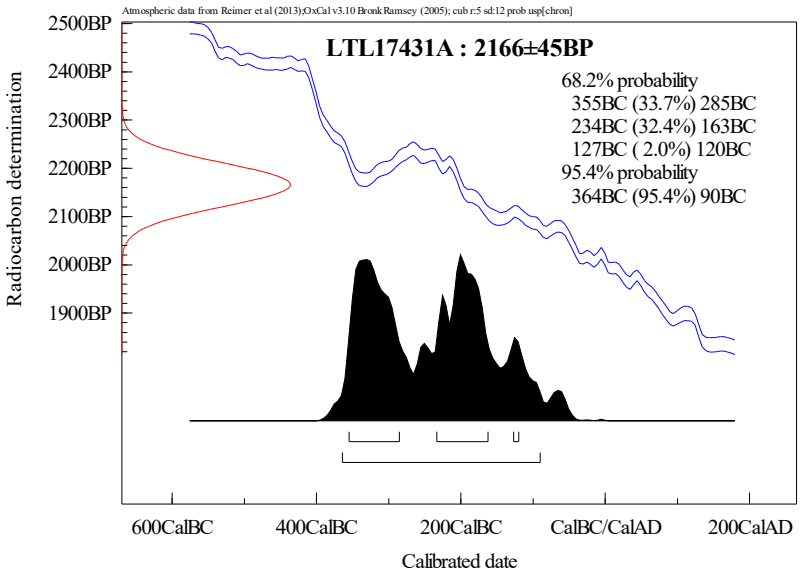


Figure 5.3.1.1b | Calibrated radiocarbon age of DA394.

5.3.2 Molecular analysis I

5.3.2.1 Authentication of the mtDNA data

Contamination from modern human DNA cannot yet be totally ruled out, and a strong logical chain of evidence is required to authenticate aDNA results (Gilbert et al., 2005). Indeed, sporadic contamination can still be observed, especially when amplifying human mtDNA with high PCR cycle numbers. In this study, it is possible to exclude contamination with a high level of confidence and attest to the authenticity of mtDNA results on the following grounds:

- *(i)* samples were collected during the archaeological excavation in a virtually modern human 'DNA-free' conditions, a circumstance that has been suggested to facilitate the discrimination between endogenous and contaminant DNA (Sampietro et al., 2006; Pruvost et al., 2007; Llamas et al., 2017);
- *(ii)* the analyses were undertaken in a dedicated aDNA laboratory of the University of Bologna, under strictly controlled conditions described in 'Ancient DNA procedures' section of Case Study I;
- *(iii)* no systematic contamination was ever observed in either the extraction or the amplification negative controls;

- *(iv)* aDNA data was considered as genuine whenever a clear sequence was reproduced in all the overlapping fragments portion of each adjacent fragment;
- *(v)* the specimens collected from Tharros were found not associated with any other part of the individuals' skeleton. Therefore, it was not possible to collect two or more samples from the same individual with the objective to replicate molecular analysis independently. Anyway, during the preparation step, all the specimens were divided and stored in two different aliquots and then analysed in different experimental batches at different times. The HVS-I mtDNA sequences were considered as authentic when they showed the same haplotype motifs.
- *(vi)* all HVS-I sequences obtained from Tharros samples showed different haplotypes from those of operators involved in this study (Table 5.3.2.1.1);
- *(vii)* the phylogenetic consistency of the haplotypes and matching haplogroup assignments of both HVR-I data and coding region SNPs were indicative of the robustness of the mtDNA typing approach presented here;
- *(viii)* the molecular behaviour of the PCR products is also in agreement with what is expected from analyses of ancient samples since the use of different primers sets to amplify DNA fragments of different lengths showed that amplification success is correlated negatively with the length of the amplicons. In fact, all the ancient samples tested with the L15996-H16401 primer pair, yielded no amplification products, indicating the absence of intact modern exogenous DNA;
- *(ix)* the HVS-I mtDNA sequences obtained made phylogenetic sense and reflected polymorphisms congruent with the geographic location under study. The haplogroup and sub-haplogroup motifs were fully represented, and no sequences showed obvious conflict with haplogroup-defining segregating sites.

Researcher	*HVS-I range*	*HVSI Haplotype based on rCRS*	*Haplogroup based on HaploGrep*	*Overall Quality **
R1	15997-16400	16223T 16325C 16362C	G1a1	100,00%
R2	15997-16400	16067T 16291T	HV1	100,00%
R3	15997-16400	16298C	HV0	100,00%
R4	15997-16400	16126C 16292T 16294T 16296T	T	83,50%

Table 5.3.2.1.1 | HVS-I motifs of the researchers who had been in contact with the ancient samples during the archaeological excavation and the laboratory work. * Overall Quality: 1 – 0.9 the haplogroup assignment is quite reliable 0.9 – 0.7 the haplogroup assignment is more accurate < 0.7 the haplogroup assignment is not very reliable.

5.3.2.2 Intra-population analysis

5.3.2.2.1 *Haplogroup assignment*

HVS-I mitochondrial consensus sequences were obtained in 10 out of the 14 ancient specimens (Table 5.3.2.2.1.1). This result represents an overall success rate of 71.42% for extraction and sequencing of mtDNA. The remaining 4 samples were excluded from subsequent analyses because yielded no amplification products (TH3 and TH18) or produced ambiguous sequence results (TH11 and TH13).

Sample	*Researchers*[1]	*rCRS position*	*HVS-I haplotype*	*SNP in coding region*	*Haplogroup*
TH1	R1, R2, R3, R4	16024-16383	16093C 16212G 16222T 16255A	7028C	H*
TH2	R1, R2, R3, R4	16024-16383	CRS	7028C	H*
TH3	R1, R2, R3, R4	16024-16383	-	-	-
TH5[2]	R1, R2, R3, R4	16024-16383	CRS	-	H
TH6	R1, R2, R3, R4	16024-16383	CRS	7028C	H*
TH7	R1, R2, R3, R4	16024-16383	16256T 16270T 16399G	7028C	H*
TH8[2]	R1, R2, R3, R4	16024-16383	CRS	-	H
TH9	R1, R2, R3, R4	16024-16383	CRS	3010A	H1
TH10[2]	R1, R2, R3, R4	16024-16383	16093C 16212G 16222T 16255A	-	H1ah2
TH11	R1, R2, R3	16024-16383	-	-	-
TH13	R1, R2, R3, R4	16024-16383	-	-	-
TH14	R1, R2, R3	16024-16383	16129A 16223T 16391A	4529T	I
TH15	R1, R2, R3	16024-16383	16129A 16223T 16391A	4529T	I
TH18	R1, R2, R3	16024-16383	-	-	-

Table 5.3.2.2.1.1 | mtDNA sequences of samples from Tharros site. mtDNA haplotypes were numbered according to the rCRS (Andrews et al., 1999). [1]For HVS-I motifs of the researchers see Table 5.3.2.1.1; [2]Haplogroup assignment using Haplogrep2 software.

Haplogroups preliminarily inferred by HVS-I motif were confirmed by the genotyping of 22 mtDNA SNPs, for which only for three samples the multiplex amplification failed (TH5, TH8 and TH10). By combining sequence and genotyping analyses the Punic samples were classified as belonging to five different mtDNA lineages: H* (TH1, TH2, TH6 and TH7), H (TH5), H1 (TH9), H1ah2 (TH10) and I (TH14 and TH15).

In this regard, it should be noted that most of the mtDNA sequences belonging to haplogroup H (hg H). The hg H clearly dominates the mitochondrial DNA (mtDNA) gene pool of Europeans (~40-45% on average) (Pereira et al., 2005; Roostalu et al., 2007), and has been a focus of attention in human genetic diversity studies for more than a decade (Achilli et al., 2004; Loogväli et al., 2004; Pereira et al., 2005; Roostalu et al., 2007; Alvarez-Iglesias et al., 2009; Behar et al., 2012, Brotherton et al., 2013). Phylogeographic studies suggest that hg H arrived in Europe from the Near East prior to the Last Glacial Maximum (~21,000 years ago, ya), and survived in glacial refugia in Southwest Europe before undergoing a postglacial

re-expansion (Soares et al., 2010; Pereira et al., 2005). Hg H lineages spread also outside of Europe and a pertinent example is found in North Africa (Cherni et al., 2009; Ottoni et al., 2010). However, it is appropriate to specify that the classification of H mtDNA samples in sub-lineages with only control region variants, such as TH5 and TH8 samples, has proven in most cases to be unreliable, due to the recurrence of some polymorphisms and the absence of diagnostic sites (Loogväli et al., 2004). In fact, complete mitochondrial genomes studies revealed that 71% of hg H polymorphic diversity is located outside the D-loop, in the coding region (van Oven et al., 2009).

The second mtDNA haplogroup found in the Punic samples is the hg I, which it is a subclade of haplogroup N1a1b and a sibling of haplogroup N1a1b1 (Olivieri et al., 2013). It is supposed to have arisen somewhere in Near East (Terreros et al., 2011; Fernandes et al., 2012) during the Last Glacial Maximum or pre-warming period (the period of gradual warming between the end of the LGM, ~19 thousand years ago - kya, and the beginning of the first main warming phase, ~15 kya) (Olivieri et al., 2013). Some subclades (I1a1, I2, I1c1, I3) show signs of the Neolithic diffusion of agriculture and pastoralism within Europe (Olivieri et al., 2013).

Recently, hg I (subclade I3) was found in a sample coming from Su Carroppu rock shelter of the Sulcis region (Sardinia, Italy) dated to 9124-7851 BC (Modi et al., 2017). In the Near East (Levant) hg I was found in a sample dated back to 8,850–8,750 years before past (yBP), while in Iran was found a sample with subclade I1c dated to 5,105 ± 35 yBP (Lazaridis et al., 2016). In Neolithic Spain hg I was found in a sample dated to 6090-5960 BC (Olivieri, et al., 2013). As mentioned above, hg I shows a strong correlation with the Indo-European migrations; in particular I1, I1a1 and I3a subclades (Olivieri et al., 2013), which have been found in Poltavka and Srubnaya cultures in Russia (Mathieson et al., 2015), among ancient Scythians (Der Sarkissian 2011), in Corded Ware and Unetice Culture burials in Saxony (Brandt et al., 2013), and in two late Neolithic individuals from Germany (Haak et al., 2015). Haplogroup I was also found among ancient Egyptian mummies excavated at the Abusir el-Meleq archaeological site in Middle Egypt, which date between the Pre-Ptolemaic/late New Kingdom to the Roman periods (Schuenemann et al., 2017) and, with a significant frequency in more recent historic grave sites (Melchior et al., 2008).

5.3.2.2.2 *Real-Time quantification*

Real-Time quantification analysis was useful to identify both the state of conservation and the sex of the samples analysed. Even though the reaction failed for TH3, TH5, TH10, TH13, TH18, due to the presumably low amount of genetic material in the specimens associated with high DNA degradation, some results were obtained for the rest of the samples (Table 5.3.2.2.2.1).

Sample	*Sample info*	*LA-large autosomal (pg/µL)*	*SA-small autosomal (pg/µL)*	*Y (pg/µL)*	*DI-Index of degradation*	*M:F*
TH1	Grave A2, S.U. 178	9.88228E-05	0.000204929	7.8528E-05	und.	0
TH2	Grave A2, S.U. 178	und.	0.000313448	0.000386794	und.	0
TH3	Grave A2, S.U. 178	-	-	-	-	-
TH5	Grave A2, S.U. 188	-	-	-	-	-
TH6	Grave A2, S.U. 188	und.	0.000162978	0.000167199	und.	0
TH7	Grave A2, S.U. 178	6.24576E-05	0.000601622	0	9.741138458	0
TH8	Grave A2, S.U. 178	und.	5.7982E-05	0	Und.	0
TH9	Grave A2, S.U. 178	und.	0.000629574	0	und.	0
TH10	Grave A2, S.U. 178	-	-	-	-	-
TH11	Grave Z, S.U. 167	-	-	-	-	-
TH13	Grave A2, S.U. 179	-	-	-	-	-
TH14	Grave Z, S.U. 167	und.	0,0001363162	0,000147505	und.	0
TH15	Grave Z, S.U. 167	und.	0,0077553727	0,013247462	und.	0
TH18	Grave A2, S.U. 178	-	-	-	-	-

Table 5.3.2.2.2.1 | DNA quantification analysis using Quantifiler® Trio kit. Abbreviation: und., undetermined.

In most of the extracts (TH1, TH2, TH6, TH7, TH8, TH9, TH14, TH15) only the small autosomal target was amplified, probably indicating a high degradation of the DNA molecules. Since for most of the analysed samples no result was obtained by the large autosomal target (LA), the degradation index (DI) was evaluated for only two samples (TH1 and TH7). The estimated values indicate the presence of a significant degradation of the extracted genetic material, suggesting that the two DNA templates are of ancient origin. Forensic and ancient DNA samples may contain mixtures of DNA from multiple individuals. The Quantifiler® Trio kit allowed to calculate the ratio of total autosomal DNA to the male-specific Y-chromosome DNA (M:F), in order to check possible contamination. The obtained results showed the absence of this kind of problem.

5.3.2.2.3 *Possible kinship relationship*

Due to the systematic clandestine excavations of the Tharros archaeological site during the 19th century, which has altered the anatomical connection of the skeletons and determined a fragmentation and incompleteness of osteological remains, it was not possible to establish the exact number of individuals buried in the A2 and Z graves using the only morphological studies. Based on the data obtained from the molecular analysis, however, it was possible to observe that:

- samples TH1 (S.U. 178, male, long bone) and TH10 (S.U. 178, femur LS) from the grave A2, showed the same HVS-I mutation motif. Sample TH1 was classified as haplogroup H*, but it was not possible to confirmed if both the samples shared the same haplogroup since the genotyping reaction failed for TH10 specimen;
- samples TH2 (S.U. 178, male, long bone), TH5 (S.U. 188, femur) and TH6 (S.U. 188, male, femur), coming from the grave A2, shared the same sequence of the rCRS. The genotyping analysis confirmed that TH2 and TH6 belong to the same paragroup H* (no result was obtained for TH5);
- samples TH8 (female, femur LS) and TH9 (female, femur RS) from the grave A2 (S.U. 178) were both female and showed the same sequence of the rCRS. TH9 was classified as haplogroup H1, but no genotyped result was obtained for TH8;
- sample TH7 (S.U. 178, female, femur RS) from the grave A2, carrying the HVS-I sequence substitutions 16256T, 16270T, 16399G and the mtDNA SNP mutation 7028C, was classified as H*;
- samples TH14 (male, tooth) and TH15 (male, tooth) coming from the S.U. 167 of the grave Z belong to the same haplogroup I, which was confirmed by the mtDNA SNPs analysis. They shared the same HVS-I motif (16129A, 16223T, 16391A).

5.3.2.3 Inter-population diversity

5.3.2.3.1 Summary and population differentiation statistics

The values of genetic and standard diversity indexes estimated for the Punic samples from Tharros are listened in Table 5.3.2.3.1.1. The same values calculated for other present-day populations are shown for a comparison (Table 5.3.2.3.1.1). As described above, samples from the necropolis of Tharros were found not associated with any other part of the human skeleton. As a result, it is possible that two or more samples within the analysed set may belong to the same individual (see 'Possible kinship relationship' section). In addition, the

fact that some individuals buried in the same grave could be maternally related might introduce a bias in the statistical analyses of the genetic data obtained. Taking into account this consideration, haplotypes observed multiple times in the same grave, in conjunction with the results of the Real-Time (sex determination) and the type of the bone analysed, were counted only once, with the hypothesis that were originated from the same or related individuals.

Populations	*Code*	n	k	Gene diversity (*h*) ± sd	MNPD (π) ± sd	Nucleot.Div. (π_N) ± sd
EUROPE						
ITALY						
Tharros, Sardinia	**TH**	**5**	**4**	**0.9000 +/- 0.1610**	**3.393136 +/- 2.080948**	**0.009425 +/- 0.006758**
Cabras, Sardinia	CB	48	35	0.9743 +/- 0.0138	5.188818 +/- 2.556029	0.016015 +/- 0.008756
Medio Campidano, Sardinia	CM	23	16	0.9130 +/- 0.0525	4.424233 +/- 2.264840	0.013655 +/- 0.007796
Oristano, Sardinia	OR	39	32	0.9865 +/- 0.0104	5.147181 +/- 2.548216	0.014298 +/- 0.007864
Olbia/Nuoro, Sardinia	OTNU	31	22	0.9591 +/- 0.0247	4.793403 +/- 2.406330	0.013315 +/- 0.007437
San Pietro, Sardinia	SP	44	33	0.9725 +/- 0.0161	4.859995 +/- 2.416004	0.013500 +/- 0.007452
Sant'Antioco, Sardinia	SA	42	29	0.9419 +/- 0.0287	4.257881 +/- 2.153625	0.011827 +/- 0.006644
Nuoro, Sardinia	NUO	51	31	0.9373 +/- 0.0254	3.887609 +/- 1.983500	0.010799 +/- 0.006114
Gallura, Sardinia	GAL	50	31	0.9371 +/- 0.0265	4.071498 +/- 2.064914	0.011310 +/- 0.006365
Trexenta, Sardinia	TX	47	31	0.9426 +/- 0.0265	4.751810 +/- 2.365636	0.013199 +/- 0.007294
Ragusa/Siracusa, Sicily	RGSR	39	28	0.9690 +/- 0.0167	5.211007 +/- 2.576252	0.014475 +/- 0.007951
Catania, Sicily	CT	37	30	0.9670 +/- 0.0224	5.784393 +/- 2.831274	0.016068 +/- 0.008740
Agrigento, Sicily	AG	42	28	0.8931 +/- 0.0457	4.200808 +/- 2.128510	0.011669 +/- 0.006566
Trapani, Sicily	TP	40	25	0.9423 +/- 0.0266	4.488651 +/- 2.257396	0.012468 +/- 0.006966
SPAIN						
Granada	GRA	121	72	0.9731 +/- 0.0083	4.381475 +/- 2.179360	0.012171 +/- 0.006703
Huelva	HUE	158	84	0.9666 +/- 0.0091	5.580611 +/- 2.694368	0.015502 +/- 0.008284
Majorca	MAJ	67	41	0.9521 +/- 0.0191	5.014945 +/- 2.467629	0.013930 +/- 0.007599
PORTUGAL						
Northern	POR_N	100	67	0.9533 +/- 0.0162	5.207707 +/- 2.541055	0.014466 +/- 0.007818
Central	POR_C	82	62	0.9768 +/- 0.0107	5.302650 +/- 2.586931	0.014730 +/- 0.007962
Southern	POR_S	59	41	0.9433 +/- 0.0246	4.934064 +/- 2.436547	0.013706 +/- 0.007507
FRANCE						
Corsica	COR	53	36	0.9586 +/- 0.0187	4.078643 +/- 2.066077	0.011330 +/- 0.006368
GREECE						
Greek from Attica	GR_A	29	25	0.9877 +/- 0.0133	4.806812 +/- 2.416895	0.013352 +/- 0.007473
Greek from Thrace	GR_T	25	20	0.9633 +/- 0.0291	5.885093 +/- 2.907982	0.016347 +/- 0.009002
Greek from Crete	GR_C	18	13	0.9477 +/- 0.0392	5.308803 +/- 2.688502	0.014747 +/- 0.008351
CYPRUS						
Cypriots	CYP	90	59	0.9793 +/- 0.0071	5.724477 +/- 2.767441	0.015901 +/- 0.008516
NORTH-AFRICA						
LIBYA						
Libyans	LYB	269	160	0.9882 +/- 0.0027	7.138372 +/- 3.358613	0.019829 +/- 0.010321
Fezzan (Al Awaynat)	ALA	111	17	0.6637 +/- 0.0493	4.359248 +/- 2.171054	0.012109 +/- 0.006678
Fezzan (Tahala)	TAH	18	6	0.7320 +/- 0.0961	5.777569 +/- 2.899807	0.016049 +/- 0.009008

Table 5.3.2.3.1.1 | Standard and genetic diversity indexes estimated for the Punic population. (continued)

Populations	*Code*	n	k	Gene diversity (*h*) ± sd	MNPD (π) ± sd	Nucleot.Div. (π_N) ± sd
MOROCCO						
Moroccans	MOR	149	62	0.9225 +/- 0.0177	4.311972 +/- 2.146429	0.011978 +/- 0.006600
TUNISIA						
Qalaat El Andalous	QEA	29	17	0.9458 +/- 0.0237	6.872243 +/- 3.329954	0.019090 +/- 0.010296
Capital Tunis	TS	51	44	0.9922 +/- 0.0062	7.932611 +/- 3.749265	0.022035 +/- 0.011557
El Alia	AL	48	27	0.9601 +/- 0.0159	6.237814 +/- 3.014329	0.017327 +/- 0.009294
Zriba	ZRI	34	18	0.9305 +/- 0.0295	7.810836 +/- 3.727464	0.021697 +/- 0.011513
Slouguia	SLA	27	19	0.9687 +/- 0.0195	6.784556 +/- 3.298592	0.018846 +/- 0.010204
Testour	TES	49	35	0.9566 +/- 0.0216	7.791223 +/- 3.690255	0.021642 +/- 0.011377
Skira	SK	20	14	0.9368 +/- 0.0427	6.451916 +/- 3.187604	0.017922 +/- 0.009889
Kesra	KES	43	30	0.9601 +/- 0.0202	7.409564 +/- 3.532597	0.020582 +/- 0.010897
ALGERIA						
Algerians	ALG	240	140	0.9671 +/- 0.0080	6.051892 +/- 2.892534	0.016811 +/- 0.008890
NEAR EAST						
PALESTINE						
Palestinian Israel	PAL	211	156	0.9950 +/- 0.0013	6.829338 +/- 3.228532	0.018970 +/- 0.009923
SYRIA						
Syrians	SYR	106	100	0.9987 +/- 0.0016	6.426845 +/- 3.067257	0.017852 +/- 0.009436
JORDAN						
Jordan	JOR	145	97	0.9784 +/- 0.0062	6.080928 +/- 2.911597	0.016891 +/- 0.008953
LIBANON						
Lebanese	LEB	363	207	0.9841 +/- 0.0032	5.440665 +/- 2.626192	0.015113 +/- 0.008069

Table 5.3.2.3.1.1 | Standard and genetic diversity indexes estimated for the Punic population.

5.3.2.3.2 *FST and Multidimensional Scaling (MDS)*

The matrix of Fst values relating to 41 contemporary populations from Europe, North Africa, and Near East with the Punic individuals from Tharros site was reported below as a heat map in order to facilitate the description of the results (Figure 5.3.2.3.2.1).

Fst values ranged from -0.00066 to 0.02228. The nearest populations to Tharros individuals are Trexenta (TX) (Fst = -0.00066) from Sardinia, Algerians (ALG) (Fst = -0.00176) and Spanish from Majorca Island (MAJ) (Fst = -0.00392). The highest genetic distance for Punic individuals is observed with Olbia/Nuoro populations from Sardinia (OT/NU) (Fst = 0.02228), Ragusa/Siracusa population from Sicily (RG/RS) (Fst = 0.02056) and Fezzan - Al Awaynat from Lybia (ALA) (Fst = 0.01226).

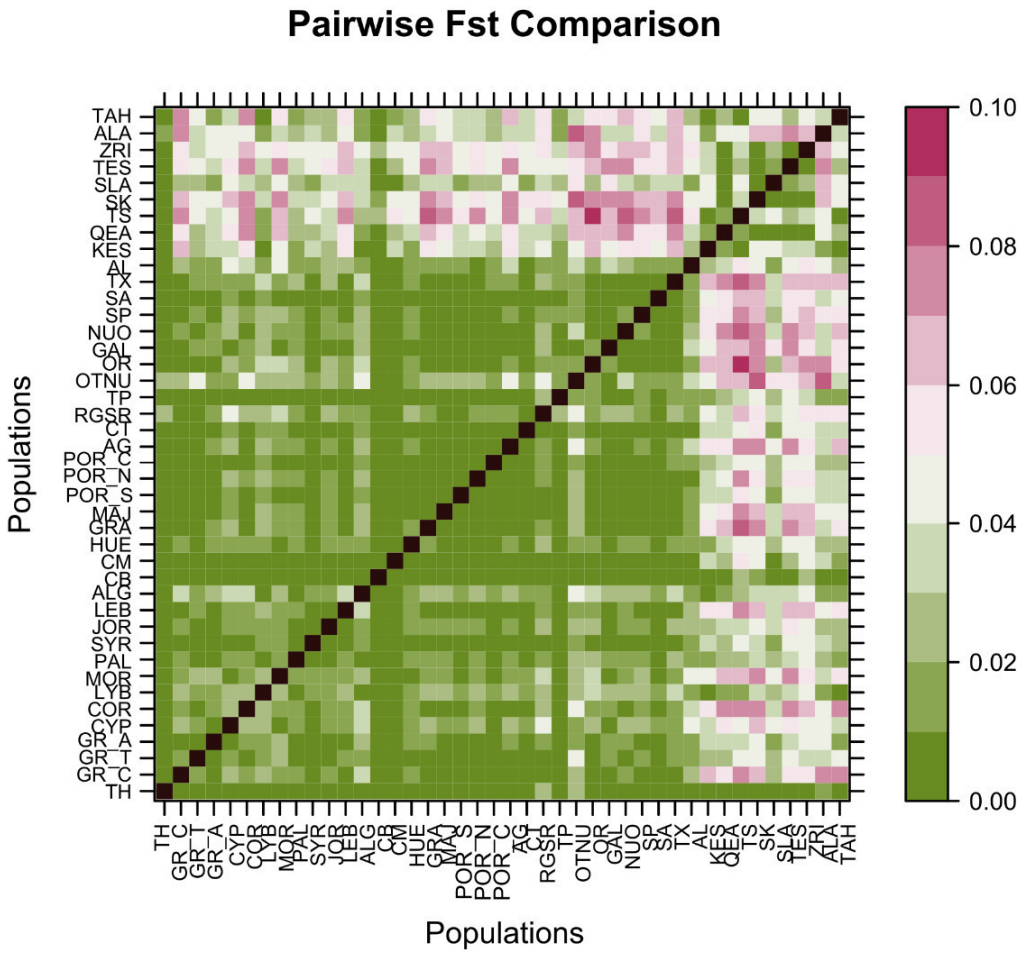


Figure 5.3.2.3.2.1 | Genetic distances (pairwise Fst) between Tharros samples (TH) and 41 present-day population of the Mediterranean Basin. Code name are reported in the Table 5.3.2.3.1.1.

A Multidimensional scaling (MDS) analysis comparing the HVS-I mitochondrial variability (np 16024-16383) of the Tharros samples and 3258 mtDNA belonging to 41 present-day Mediterranean populations was carried out to provide a two-dimensional plot of the Fst genetic distances matrix (Figure 5.3.2.3.2.2). All North African samples are deployed on the left side of the plot, except the Moroccan population, which is displayed close to Near Eastern and the South-Western and Eastern European populations on the opposite side. Punics from Tharros site occupy an intermediated position between North African populations and the bulk of Near Eastern and South-European samples. Instead, all the Sardinian populations are rather distant from the old samples here analysed. It is also interesting to note the proximity of the present-day inhabitants of Cabras with the ancient Punics individuals thus suggesting a certain degree of genetic continuity, between Punics from Tharros and the population that nowadays living in the same territory.

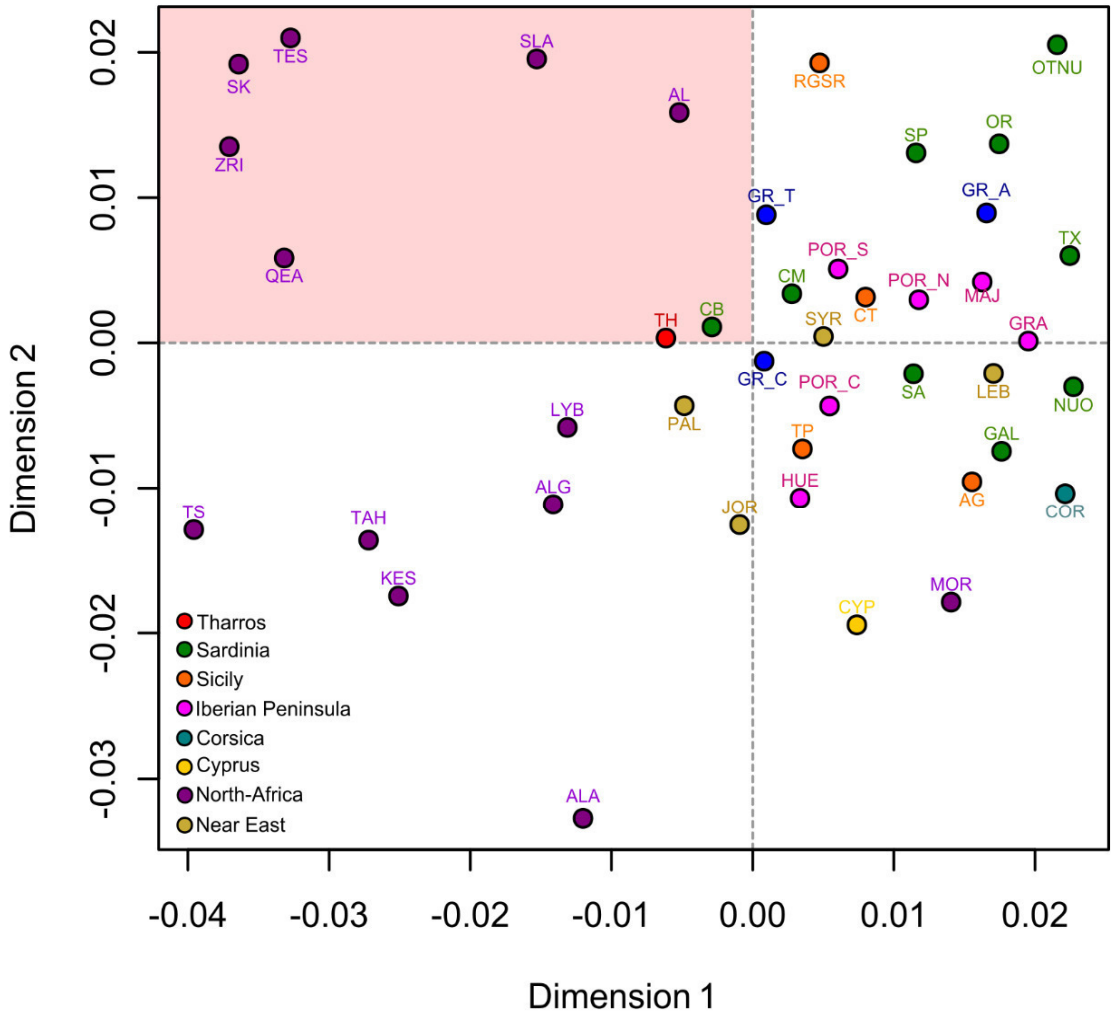


Figure 5.3.2.3.2.2 | Two-dimensional MDS plot based on Slatkin's Fst showing genetic affinities among samples from Tharros (TH) and 41 present-day population of the Mediterranean Basin.

To better clarify the genetic affinities among the populations included in the dataset, a PCA analysis was performed. The scatter-plot of the first two PCs computed on mtDNA HVS-I data, attested the overall resemblance between genetic relationships and geographic distribution of the examined populations. The PCA (Figure 5.3.2.3.2.3) shows that the cluster of the North African populations are disposed on left side of the plot, and the group of the South-European populations (i.e. Portugal, Spain, Sicily, Greek and Corsica samples) are located at the centre of the graph, just below the group constituted by Cyprus and Near Eastern samples. All the Sardinians cluster together on the right side of the scatter-plot, outside to the variability of the other South-European populations. This outcome is in line whit the several population-genetics studies so far carried out, which showed that the Sardinian population is one of the main European genetic outliers with a unusually high levels of internal diversity (Zei et al., 2003; Cavalli-Sforza et al., 2004; Pala et al., 2009; Francalacci

et al., 2013; Sidore et al., 2015; Olivieri et al., 2017). Ancient Punic samples appeared to be genetically closer to the North African populations, especially with Tunisia, rather than with Sardinian.

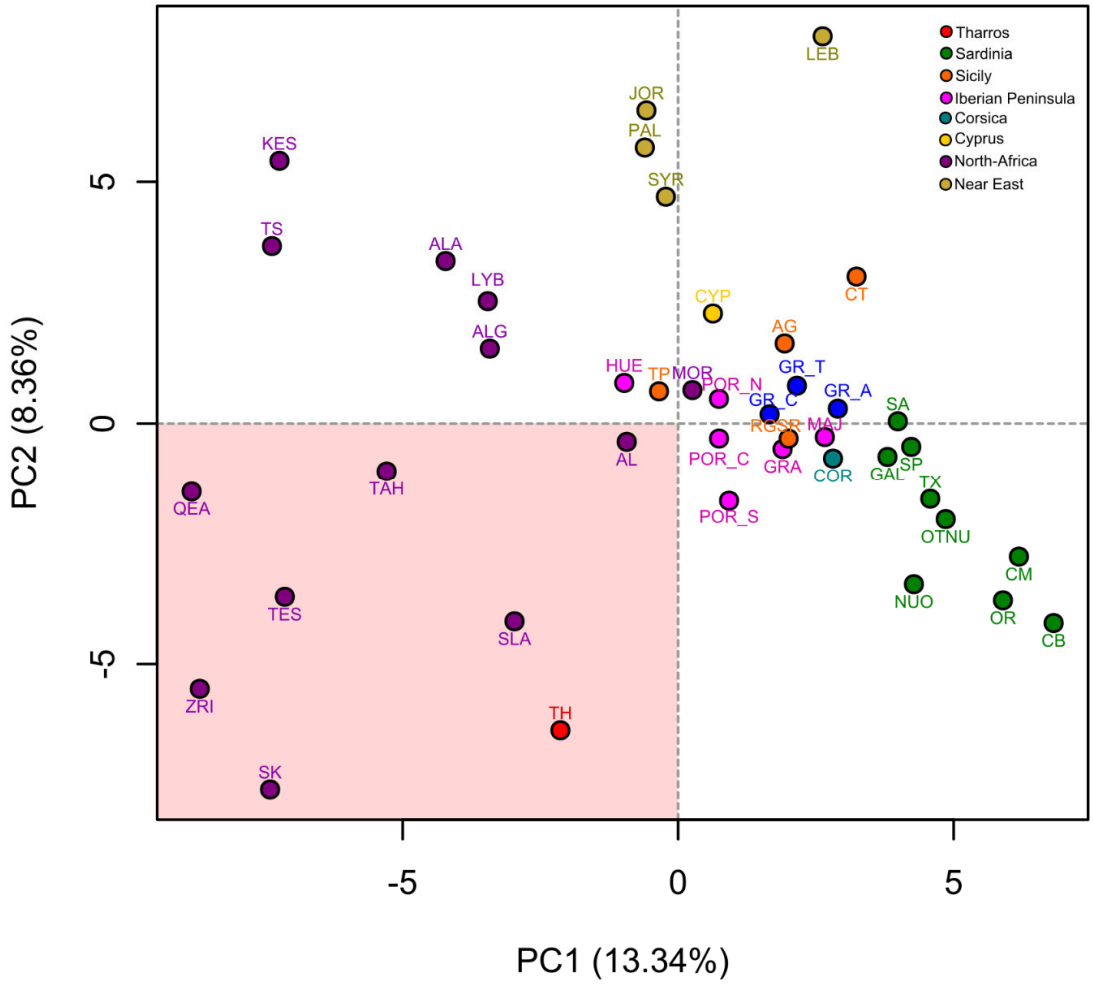


Figure 5.3.2.3.2.3 | PCA plot showing genetic affinities among samples from Tharros (TH) and 41 present-day population of the Mediterranean Basin.

5.3.3 Molecular analysis II

5.3.3.1 Endogenous nuclear DNA content

From all 14 shotgun-sequenced libraries, were obtained a total of 293 697 606 reads. Statistics of the read data processing is shown in (Table 5.3.3.1.1).

Sample	*Total*	*Trimmed*	*Unique*	*Rmdup*	*Clon (%)*	*Endo (%)*
DA390	5482545	5477739	26090	25589	3.09404423	0.5
DA391	12452706	12443940	1919	1897	3.53505566	0
DA392	34591161	33988248	160480	160377	0.57515034	0.5
DA393	20963082	19544379	7393725	7348810	0.607474581	37.8
DA394	19764288	19119430	607290	607109	0.149031	3.2
DA395	16148650	15817906	793	780	0	0
DA396	17889432	17490634	1168	1090	0	0
DA397	19310954	18455375	770	743	0	0
DA398	22437333	21485586	5679	5676	0.289709	0
DA399	21867998	20856695	91292	91200	0.571985	0.4
DA400	26425723	24709452	1711	1698	0	0
DA401	24679755	23132837	35629	35608	0.363797	0.2
DA402	26583298	25322174	240714	240583	0.370901	0.9
DA403	25100681	24796538	31577	31566	0.218341	0.1

Table 5.3.3.1.1 | Shotgun sequencing of 14 ancient Punic samples: *Total* is the total number of DNA reads per library. *Trimmed*, is number sequences passing quality and length filtering. *Unique* is the number of sequences mapping uniquely to the human reference genome. *Rmdup* is the same number but with all duplicate sequences removed. *Clon %* is the proportion of identical reads (clones) in this human DNA fraction. *Endo %* is the proportion of sequences after trimming that could be identified as human.

Unfortunately, only 1 sample out of the 14 samples tested results to have enough human DNA useful for carrying out the subsequent population analyses (Figure 5.3.3.1.2): this is the sample DA393 (petrous bone) coming from the grave L of the Tharros site. In fact, owing to high levels of endogenous DNA, the inner part of the petrous bone is currently recognized as the optimal substrate for aDNA research. Given its extremely high density, the petrous bone has been shown to yield 4- to 16-fold more endogenous DNA than teeth, and up to 183-fold more than other bones (Gamba et al., 2014).

Several visual factors can indicate good macroscopic preservation, and high chances of endogenous DNA survival. For example, fresh-looking compact bones or bone fragments with smooth and intact surfaces are indicators for good macroscopic preservation. Moreover, environmental conditions (e.g. temperature, moisture, and pH) in combination with time since death (post-mortem interval - PMI) are thought to be the primary factors influencing DNA degradation, but the relative effects of environment and time appear to be strongly situation dependent, leading to claims that the rate of DNA degradation cannot be predicted. In this case, the low amount of endogenous DNA observed in Tharros and Lilybaeum specimens (intact samples with smooth surface and fresh-looking, see Figure 5.2.5.1.1) might be explained by the relatively wary climatic conditions in both the Islands, which are not optimal for DNA preservation.

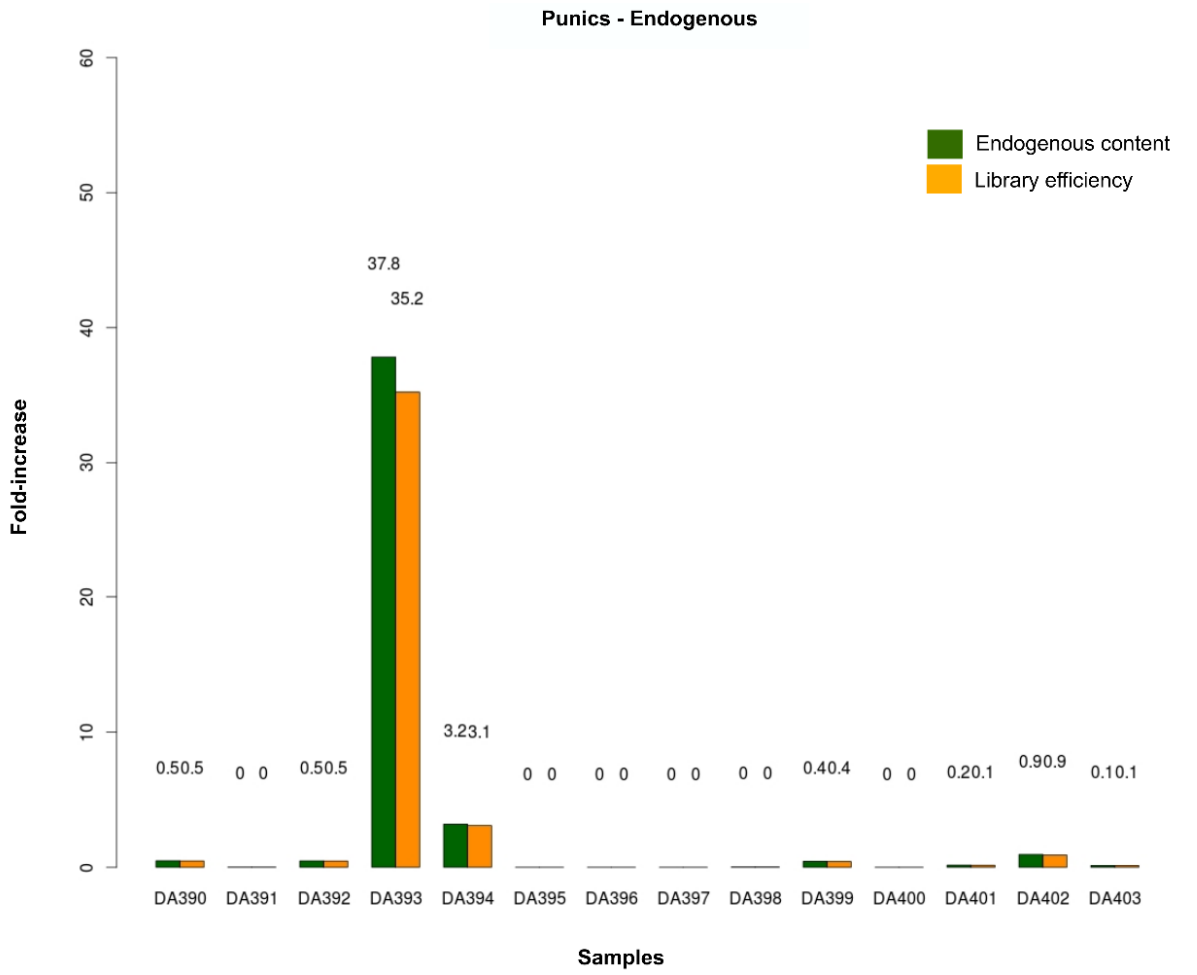


Figure 5.3.3.1.2 | Amount of endogenous content in the Punic data obtained. Green bars represent the increase in endogenous content, and yellow bars represents the increase in library efficiency.

5.3.3.2 Estimation of the contamination and authentication of data

In order to assess aDNA authenticity, the aDNA damage patterns (Figure 5.3.3.2.1a, 5.3.3.2.1b) and the mtDNA contamination (Figure 5.3.3.2.2) were estimated with mapDamage (Ginolhac et al. 2011; Jònsson et al., 2013) and contamMix (Fu et al., 2013). The results demonstrate that the sequence data of DA393 individual is endogenous and minimally contaminated.

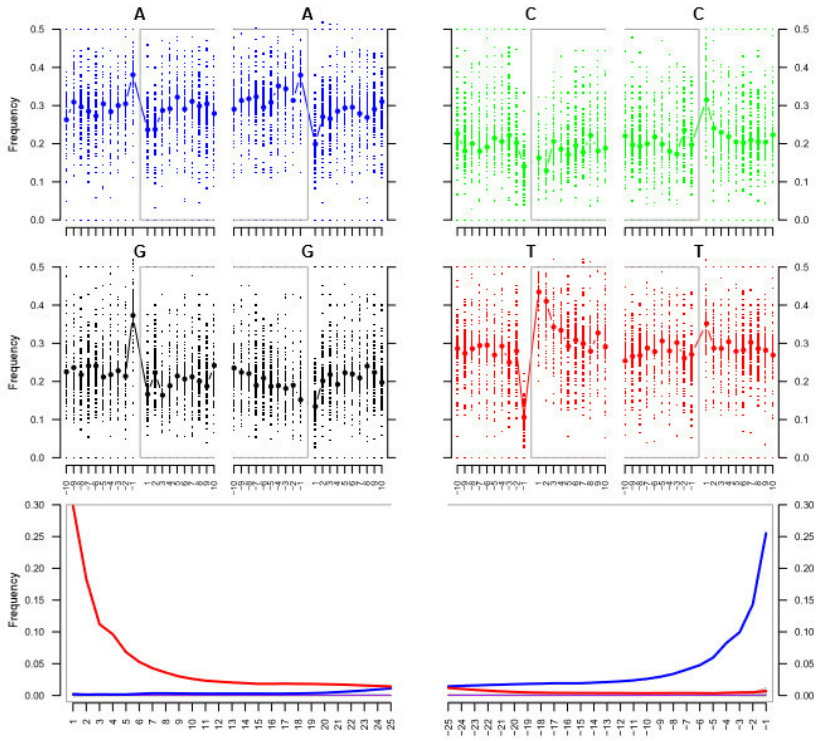


Figure 5.3.3.2.1a | mapDamage results for DA393 sample. The four upper mini-plots show the base frequency outside and in the read (the open grey box corresponds to the read). The bottom plots are the positions' specific substitutions from the 5" (left) and the 3" end (right).

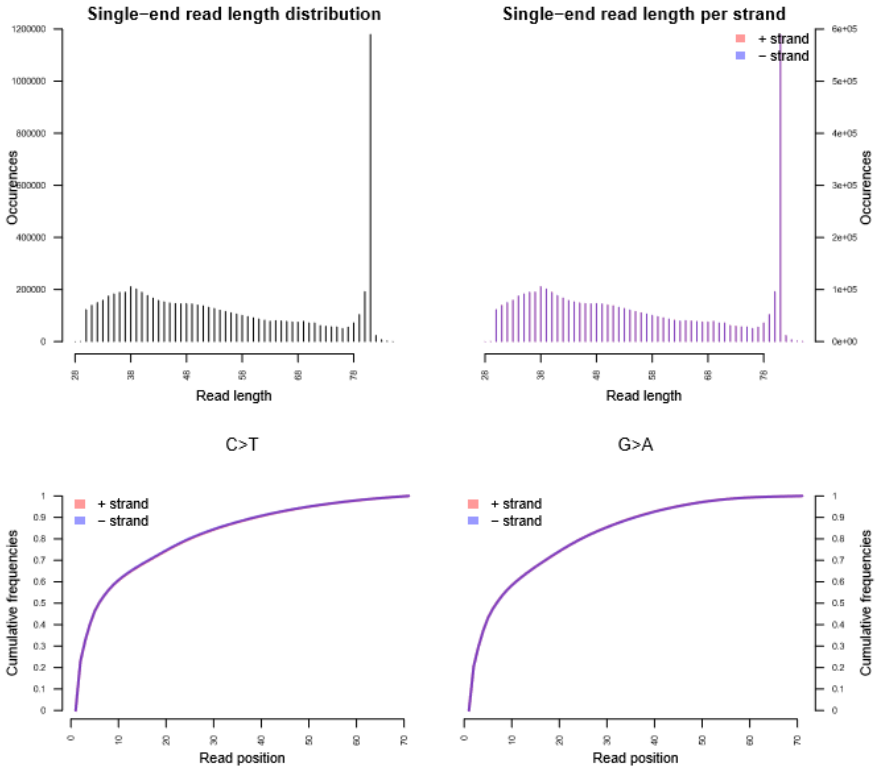


Figure 5.3.3.2.1b | mapDamage results for DA393 sample. The upper two plots are histograms of the read lengths. The lower two plots are the empirical cumulative frequency of C->T and G->A misincorporations.

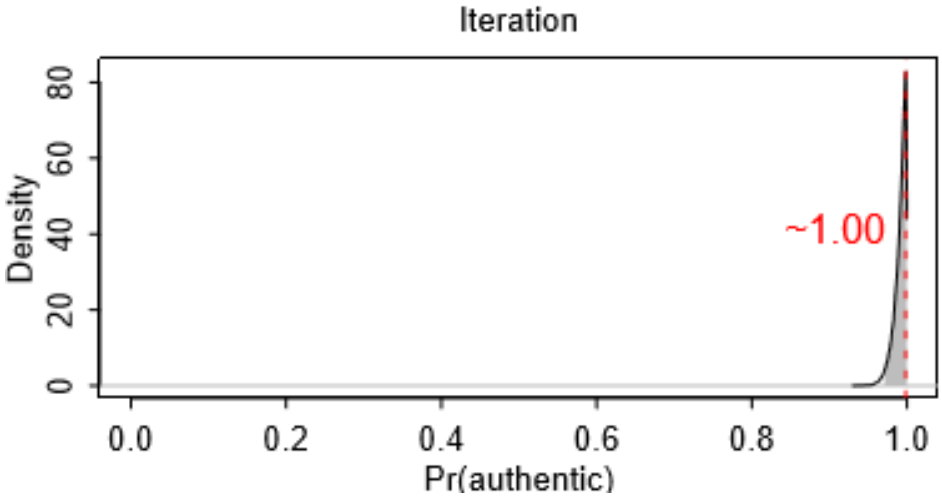


Figure 5.3.3.2.2 | Contamination pattern with contamMix (CI 3%).

5.3.3.3 Population genetic analyses

Principal component analysis (PCA) was carried out in order to shed light on the genetic history of the Punic individual DA393 from Tharros, resolving its relationship to the ancient populations and assessing its genetic contribution to present-day population. PCA was etimated from 82 modern-day populations (Figure 5.3.3.3.1) and projected the Punic individual along 390 previously published ancient individuals (grouped on the basis of archaeological culture, chronology, and genetic clustering) onto the first two principal component (Figure 5.3.3.3.2).

In accordance with its geographical location, the Punic sample (Tharros) attested its relatedness to the present-day Sardinian population. As concerns the ancient individuals included in the dataset, the Tharros sample is placed along the cline of Anatolian/European Early Neolithic and European Middle Neolithic/Chalcolithic genetic variation, occupying an intermediate, but decentralised position between these two groups. The relative closeness between the Punic sample and the present-day Sardinian suggests a certain degree of genetic continuity, probably due to the isolated position of the island. Consistently with the most recent literature (Gunther and Jakobsson, 2016), the PCA plot (Figure 5.3.3.3.2) confirms the genetic affinity of Anatolian and Early/Middle-Neolithic Europeans to modern Southern European populations, and particularly to present-day Sardinians, who are isolated from the rest of modern populations. The inhabitants of Sardinia have a distinct genetic heritage that sets them apart from other Europeans. Their peculiar history and isolation from the rest of the continent may explain why these people's genetic signatures are unique.

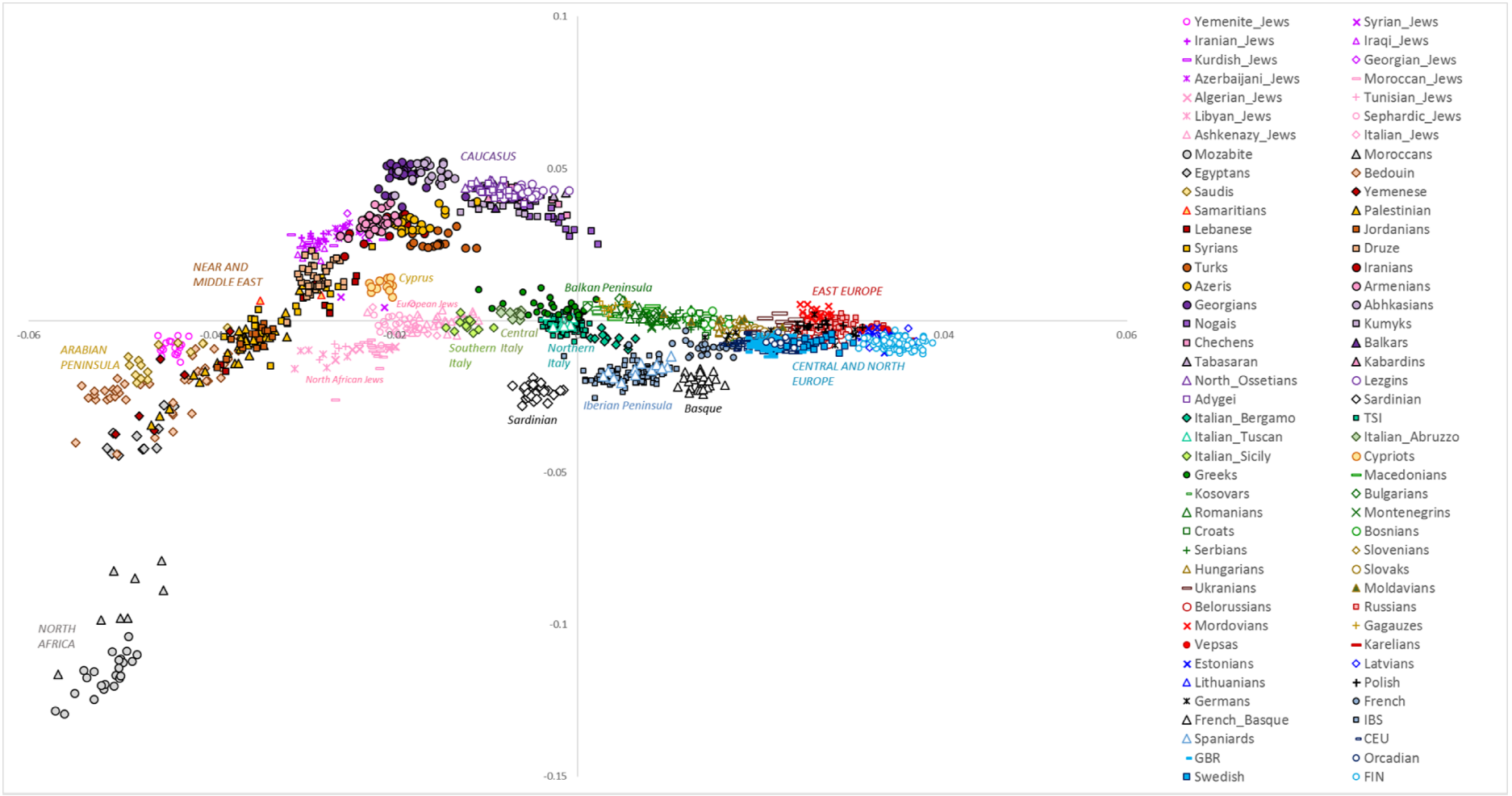


Figure 5.3.3.3.1 | PCA analysis calculated from Modern Reference Dataset.

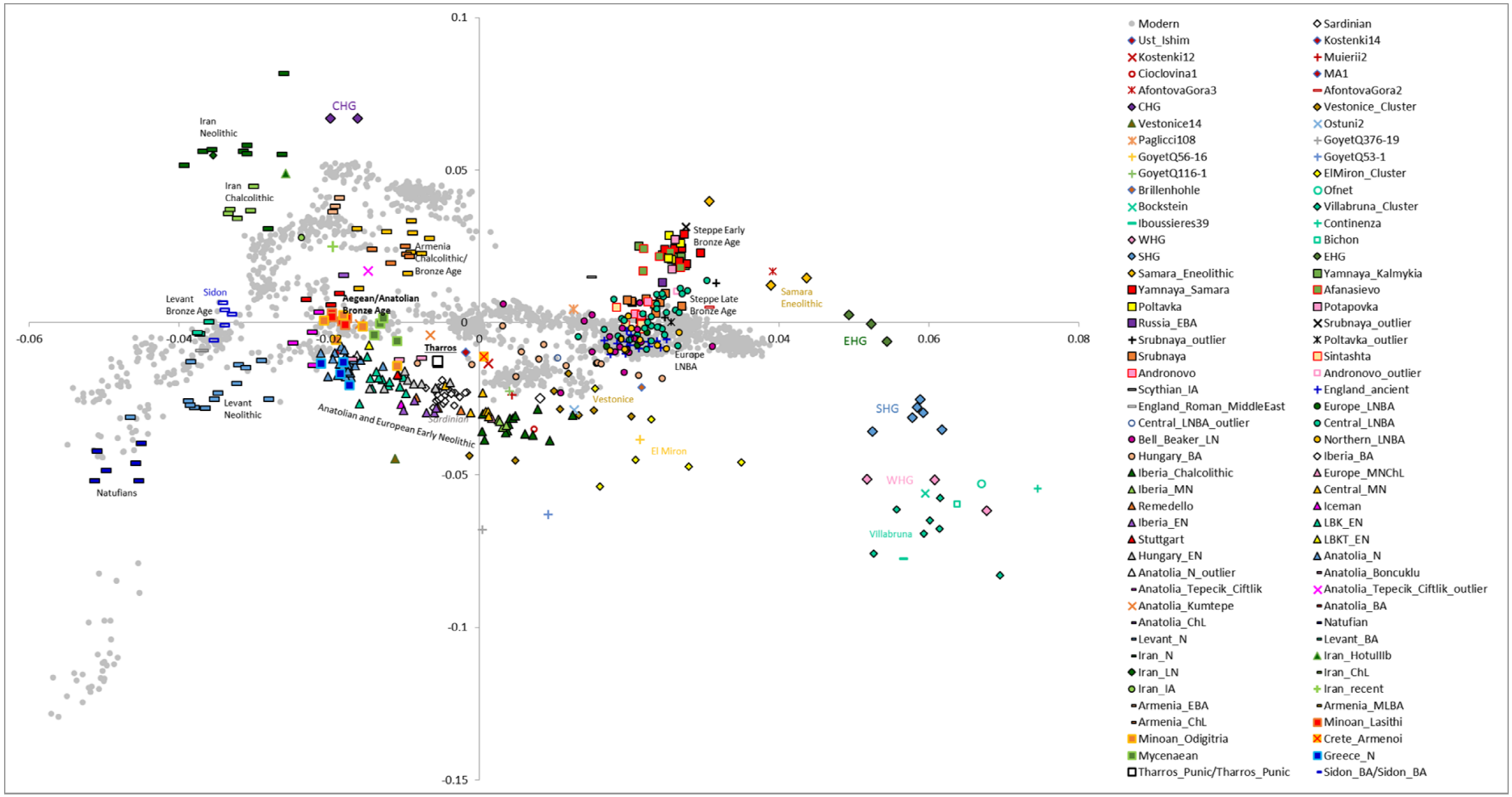


Figure 5.3.3.3.2 | PCA on contemporary populations (grey points) onto which ancient individuals are projected from this study (DA393) and previous studies (coloured shapes). Punic sample is reported as white square labelled with 'Tharros' name. Present-day Sardinian are reported as white rhombuses.

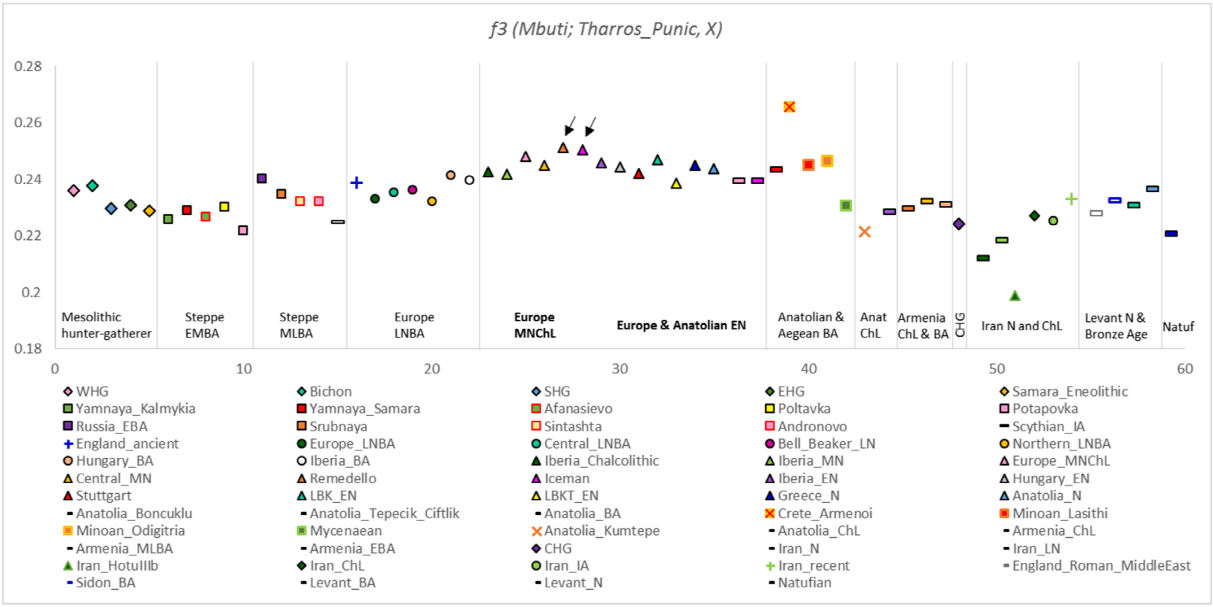


Figure 5.3.3.3.3 | Shared drift with ancient individuals using outgroup f3 statistic

Consistent to the PCA, outgroup f_3 statistic (Figure 5.3.3.3.3) confirmed the affinities of Tharros sample with the European/Anatolian Early Neolithic and European Middle Neolithic/Chalcolithic cluster, as well as Anatolian and Aegean Bronze Age groups. In particular, the Tharros sample shows higher affinity with the so-called Tyrolean Iceman from Alps (5,350-5,100 BC) (Keller et al., 2012) and the Remedello individuals from Chalcolithic Northern Italy (3,400-2,800 BC) (Allentoft et al., 2015), as well as with the Armenoi sample (Creta_Armenoi) from West Crete (1,370-1,340 BC) (Lazaridis et al., 2017), for which, however, the lack of high quality data should be taken into account.

5.4 Conclusion and Future Objectives

As known, both in ancient historical sources and in modern scholarship, the terms 'Phoenicians' and 'Punics' are merely conventional labels and are of little if any importance in relation with the complex issue of defining the ethnic identity of peoples that since the early Iron Age played a key role in the high mobility phenomena characterizing the Mediterranean area.

The ethnic composition of the crews of trade vessels, the Homer's 'black ships' sailing the *mare nostrum* with frequent stopovers and crew changes along the Aegean and centre-Western coasts, is an unresolved issue both from a historical and an archaeological point of view. As a result, the interpretation of contacts and interrelations between early 'Phoenician'

colonists and indigenous populations of North Africa, Spain, Malta, Pantelleria, Sicily and Sardinia, involved in different periods and modes, first by the Levantines and then by the Carthaginians, in the establishment of colonial foundations or trade and production enclaves, remains highly problematic. The currently available archaeological evidence offers a varied relationships picture that in some instances suggests integration or actual 'interbreeding' phenomena (such as in the material culture of some North African, Iberian or Sardinian settlements linked to productive activities that involved advanced technology transfer from Eastern "specialists" to indigenous communities), in other cases the archaeological record points to a colonial settlement system that established urban centres whose cultural identity appears to be an autarchic development, that is extraneous to forms of deep interrelationship with indigenous peoples. In the central Mediterranean area, the 'Phoenician' and 'Punic' Sardinia and Sicily offer a privileged perspective on this phenomenon.

In order to contribute to the reconstruction of the 'Phoenician' and 'Punic' settlement in the Central-Western Mediterranean area, several human remains were collected from the archaeological sites of Tharros (N = 25) (OR, Sardinia) and Lilybaeum (N = 3) (TP, Sicily) with the aim to analysed aDNA and to enlarge knowledge on Phoenician and Punic migration flows and relationship in the Mediterranean basin.

The analysis of the HVS-I still represents an evaluable approach to study mitochondrial DNA genetic diversity (see Methodological approach I). In fact, the high copy number, small genome size, maternal inheritance and the presence of a fast-evolving region, make mtDNA the selected molecule for many genetic applications, ranging from forensic investigations, genealogical purposes and population genetic studies. HVS-I mtDNA consensus sequences were obtained in 10 out of the 14 ancient specimens from southern necropolis of Tharros (Sardinia), which represents an overall success rate of 71.42% for extraction and sequencing of mtDNA. By combining sequence and genotyping analyses on the coding region (22 SNPs) the Punic samples were classified as belonging to five different mtDNA lineages: H* (n = 4), H (n = 1), H1 (n = 1), H1ah2 (n = 1) and I (n = 2). Based on the inter-population analyses, carried out by comparing Punic samples with several present-day populations of the Mediterranean basin, the Tharros samples seem to be genetically closer to the current inhabitants of the North Africa area, especially from Tunisia, rather than to modern populations of Sardinia.

Even in its earliest phases the town of Tharros, whose foundation took place between the 8th and the 7th century BC on the extreme end of the Sinis peninsula, was a cosmopolitan

centre and a port crucial to the circulation of goods and men from and to the North Africa, the Balearic Islands, Iberia and Cyprus, as it has been demonstrated by the current available archaeological and epigraphic evidence; it then became the administrative capital of Carthage and probably terminal for the arrival of North African workforce in the Punic and Roman period. Regarding the issue, still open and debated, about the dynamics of integration between Proto-Sardinian and Phoenician-Punic communities, the data provided by these preliminary analyses, do not allow us to conclude that the admixture between the two populations does not correspond to a generalized phenomenon. Although preliminary and being aware that the matrilineal DNA perspective may only reveal a part of the population history, the results obtained from the HVS-I analysis coming from grave A2 and Z, seems to highlight in the Tharros population, at least with regard to the historical period of Carthaginian control in Sardinia, a greater incidence of the North African genetic component than the autochthonous one, which is in concordance with the recently archaeological evidence (Fariselli, 2017).

With the advent of new technologies (NGS) applied to the aDNA field, it is now possible to obtain a huge amount of high-resolution genomic information (methodological approach II). So far, whole genomes have been sequenced from ancient anatomically modern humans, archaic hominins, ancient pathogens and megafaunal species. These analyses revealed important functional and phenotypic information, as well as unexpected adaptation, migration, and admixture patterns. In this study, the approach carried out started from the aDNA extraction from a total of 14 specimens, the construction of the libraries and the shotgun sequencing on NGS platform. Only 1 sample out of 14, results to have enough human DNA: this is the sample DA393 (from petrous bone), dated to the 382 ± 153 BC, coming from the hypogea grave L in the southern necropolis of Tharros. The low amount of endogenous DNA observed in Tharros and Lilybaeum specimens might be explained by the relatively wary climatic conditions in both the Islands, which are not optimal for DNA preservation. In fact, both the amount of degradation of aDNA and the amount of damage are influenced by the environmental circumstances of the sample. As known that cold and dry conditions favour the DNA preservation, while hot and humid conditions do not.

By comparing the DA393 genome with ancient and modern reference populations, the present-day Sardinians, as well as ancient individuals from the European/Anatolian Early Neolithic and European Middle Neolithic/Chalcolithic appear to be those genetically closer. In particular, the outgroup f_3 statistical analysis revealed high level of affinity with the Tyrolean Iceman (5,350-5,100 BC) (Keller et al., 2012) and the Remedello individual (3,400-

2,800 BC) (Allentoft et al., 2015) from the North Italy. From the PCA analysis, DA393 appear to be genetically distant to all the clusters of the present-day Middle/Near East and North Africa populations, as well as with the five ancient Canaanites (3,750-3,650 BC) (Haber et al., 2017) from Sidon in Lebanon (historic Phoenician homeland). The fact that DA393 seems to be genetically related to the modern Sardinian, rather than to the modern Phoenician-influenced populations of the Mediterranean basin, such as Middle/Near East, North Africa, Sicily, Spain, and Cyprus, suggests that the ancient samples from Tharros probably was a 'Sardinian autochthonous'. Sardinia is known to have remained unconnected to the Italian mainland, even when sea level was at its lowest during the Last Glacial Maximum, which reached its peak about 21,500 ya. As such, their interactions with other Europeans have remained limited over thousands of years. Previous researches carried out on the genomes of Sardinians have revealed that they show high levels of nuclear genome similarity with European Neolithic farmers as well as with the Tyrolean Iceman.

In order to better clarify this situation, it would be appropriate and fruitful to broaden the dataset with a higher number of ancient genomes of the same culture, territory and historical period to have a better contextualization in time and space. This study is one of the first investigation that deals with the issue of the 'Phoenician' identity through the genetic study of ancient remains, as only two previous studies have been conducted on DNA analyses, one on modern population (Zallua et al., 2008) and the second limited to the study of the mitochondrial DNA genome of an individual uncovered in the 6th BC Carthaginian necropolis of Byrsa (Matisoo-Smith et al., 2016).

References:

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, et al. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am. J. Hum. Genet. 75:910–8.

Acquaro E, Del Vais C, Fariselli AC, 2006 (edd.) Beni culturali e antichità puniche. La necropoli meridionale di Tharros. Tharrhica – I (= Biblioteca di Byrsa, 4), La Spezia.

Acquaro E, Mezzolani, A., 1996. Tharros, Itinerari. Libreria dello stato. Istituto poligrafico e zecca dello stato, Roma.

Allentoft ME, Sikora M, Sjögren K-G, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, et al. 2015. Population genomics of Bronze Age Eurasia. Nature 522:167–172.

Alvarez-Iglesias V, Mosquera-Miguel A, Cerezo M, Quintáns B, Zarrabeitia MT, Cuscó I, Lareu MV, García O, Pérez-Jurado L, Carracedo A, et al. 2009. New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. Macaulay V, editor. PLoS One 4:e5112.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–65.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23:147.

Aubet ME. 2009. Tiro y las colonias fenicias de Occidente.

Bartoloni P, 1981. Contributo alla cronologia delle necropoli fenicie e puniche di Sardegna, in «RStFen» 9 suppl., 1981, pp. 13-29

Becthold B, 1999. La necropoli di Lylibaeum, Trapani.

Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, Yudkovsky G, Rosengarten D, Pereira L, Amorim A, et al. 2008. Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. PLoS One 3:e2062.

Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli E-L, Silva NM, Kivisild T, Torroni A, Villems R, Darwin C, et al. 2012. A Copernican reassessment of the human mitochondrial DNA tree from its root. Am. J. Hum. Genet. 90:675–84.

Bekada A, Fregel R, Cabrera VM, Larruga JM, Pestano J, Benhamamouch S, González AM. 2013. Introducing the Algerian Mitochondrial DNA and Y-Chromosome Profiles into the North African Landscape. Pereira LMSM, editor. PLoS One 8:e56775.

Bertoncini S, Blanco-Rojo R, Baeza C, Arroyo-Pardo E, Vaquero MP, López-Parra AM. 2011. A Novel SNaPshot Assay to Detect Genetic Mutations Related to Iron Metabolism. Genet. Test. Mol. Biomarkers 15:173–179.

Boattini A, Martinez-Cruz B, Sarno S, Harmant C, Useli A, Sanz P, Yang-Yao D, Manry J, Ciani G, Luiselli D, et al. 2013. Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata. PLoS One 8:e65441.

Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, Fajkošová Z, Powell A, Burger J. 2013. 2000 Years of Parallel Societies in Stone Age Central Europe. Science (80). 342.

Bollongino R, Tresset A, Vigne J-D. 2008. Environment and excavation: Pre-lab impacts on ancient DNA analyses. Comptes Rendus Palevol 7:91–98.

Bondì SF, Botto M, Garbati G, Oggiano I. 2009. Fenici e Cartaginesi. Una civiltà mediterranea.

Bonnet C, 1993. La légende de Phoinix à Tyr, SPh. N°I-II, Leveun, pp.113-123.

Bosch E, Calafell F, Gonzalez-Neira A, Flaiz C, Mateu E, Scheil H-G, Huckenbeck W, Efremovska L, Mikerezi I, Xirotiris N, et al. 2006. Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann. Hum. Genet. 70:459–487.

Brandt G, Haak W, Adler CJ, Roth C, Szecsenyi-Nagy A, Karimnia S, Moller-Rieker S, Meller H, Ganslmeier R, Friederich S, et al. 2013. Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. Science (80-.). 342:257–261.

Brisighelli F, Álvarez-Iglesias V, Fondevila M, Blanco-Verea A, Carracedo A, Pascali VL, Capelli C, Salas A. 2012. Uniparental markers of contemporary Italian population reveals details on its pre-Roman heritage. Caramelli D, editor. PLoS One 7:e50794.

Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, Richards SM, Sarkissian C Der, Ganslmeier R, Friederich S, et al. 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat. Commun. 4:1764.

Brown TA, Nelson DE, Vogel JS, Southon JR. 1988. Improved Collagen Extraction by Modified Longin Method. Radiocarbon 30:171–177.

Caramelli D, Lalueza-Fox C, Vernesi C, Lari M, Casoli A, Mallegni F, Chiarelli B, Dupanloup I, Bertranpetit J, Barbujani G, et al. 2003. Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. Proc. Natl. Acad. Sci. 100:6593–6597.

Caruso E, 2000. Documenti e problemi di topografia storica delle città fenicio-puniche dell Sicilia occidentale: la necropoli e il tofet di Lilibeo (Marsala), in Terze giornate internazionali di studi sull'area elima, Pisa-Gibellina, 217-261.

Caruso E, 2008. Lilibeo Punics e romana: storia e topografia, in Caruso E and Spanò Giammellaro A (edd.), Lilibeo e il suo territorio, 73-76.

Cherni L, Fernandes V, Pereira JB, Costa MD, Goios A, Frigi S, Yacoubi-Loueslati B, Amor M Ben, Slama A, Amorim A, et al. 2009. Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. Am. J. Phys. Anthropol. 139:253–260.

Cooper A, Poinar HN. 2000. Ancient DNA: do it right or not at all. Science 289:1139.

Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga J-L, et al. 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. 110:15758–15763.

Dabney J, Meyer M, Pääbo S. 2013b. Ancient DNA damage. Cold Spring Harb. Perspect. Biol. 5.

Damgaard PB, Margaryan A, Schroeder H, Orlando L, Willerslev E, Allentoft ME. 2015. Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. 5:11184.

Del Vais C, 2006. Per un recupero della necropoli meridionale di Tharros: alcune note sugli scavi ottocenteschi, in: Acquaro E, Del Vais C, Fariselli AC (Eds.), Beni Culturali e Antichità Puniche. La Necropoli Meridionale Di Tharros. Tharrhica - I. Agorà Edizioni, La Spezia, pp. 7–41.

Del Vais C, 2011. Cippi e altarini dalla necropoli meridionale di Tharros. Byrsa Arte Cult. E Archeol. Mediterr. Punico 19–20, 35–60.

Del Vais C, 2013. Nuove ricerche nella necropoli settentrionale di Tharros (campagne 2010-2011): l'Area A, in ArcheoArte. Rivista elettronica di Archeologia e Arte 2, 333-34.

Del Vais C, 2013. Stele, cippi e altarini dalle necropoli puniche di Tharros, Biblioteca di Byrsa. Agorà & Co., Lugano.

Del Vais C, Fariselli AC, 2010. Tipi tombali e pratiche funerarie nella necropoli settentrionale di Tharros (San Giovanni di Sinis, Cabras -OR). Ocnus Quad. Della Scuola Spec. Beni Archeol. 18, 9–22.

Del Vais C, Fariselli AC, 2010a. Nuovi scavi nella necropoli settentrionale di Tharros (Loc. S. Giovanni di Sinis, Cabras - OR), in ArcheoArte. Rivista elettronica di Archeologia e Arte 1, 305-06.

Del Vais C, Fariselli AC, 2010b. Nuove ricerche nella necropoli settentrionale di Tharros (campagne 2010-2011): l'Area B, in ArcheoArte. Rivista elettronica di Archeologia e Arte 2, 335-36.

Del Vais C, Fariselli AC, 2012. La necropoli settentrionale di Tharros: nuovi scavi e prospettive di ricerca (campagna 2009), in: Ricerca e Confronti 2010. Atti Delle Giornate Di Studio Di Archeologia e Storia Dell'arte a 20 Anni Dall'istituzione Del Dipartimento Di Scienze Archeologiche e Storico-Artistiche Dell'Università Degli Studi Di Cagliari (Cagliari, 1-5 Marzo 2010), ArcheoArte. Rivista Elettronica Di Archeologia e Arte. Supplemento 1. pp. 261–283.

Del Vais C, Usai E, 2013. Nuove ricerche nella necropoli di Othoca (loc. Santa Severa, Santa Giusta-OR) (campagne 2010-2011). ArcheoArte Riv. Elettronica Archeol. E Arte 2.

Depalmas, A., Melis, R.T., 2010. The Nuragic people: their settlements, economic activities and use of the land, Sardinia, Italy. In: Martini, I.P., Chesworth, W. (Eds.), Landscapes and Societies. Springer Science and Business Media B.V, Dordrecht Heidelberg London New York, 177–186.

DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43:491–8.

Di Salvo R, 2004. Antropologia e paleopatologia dei gruppi umani di età fenicio-punica della Sicilia Occ, in Pràts G (ed.), El mundo Funerario, Alicante, 253-258.

Di Stefano CA, 1993. Lilibeo punica, Marsala.

Excoffier L, Lischer H. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567.

Fadhlaoui-Zid K, Martinez-Cruz B, Khodjet-el-khil H, Mendizabal I, Benammar-Elgaaied A, Comas D. 2011. Genetic structure of Tunisian ethnic groups revealed by paternal lineages. Am. J. Phys. Anthropol. 146:271–280.

Falchi A, Giovannoni L, Calo CM, Piras IS, Moral P, Paoli G, Vona G, Varesi L. 2006. Genetic history of some western Mediterranean human isolates through mtDNA HVR1 polymorphisms. J. Hum. Genet. 51:9–14.

Fariselli AC, 2006. Il "paesaggio"funerario: tipologia tombale e rituali, in: Acquaro, E., Del Vais, C., Fariselli, A.C. (Eds.), Beni Culturali e Antichità Puniche. La Necropoli Meridionale Di Tharros. Tharrhica - I. Agorà Edizioni, La Spezia, pp. 303–369.

Fariselli AC, 2008. Tipologie tombali e rituali funerari a Tharros, tra Africa e Sardegna, in L'Africa Romana. Le ricchezze dell'Africa. Risorse, produzioni, scambi. Atti del XVII convegno di studio (Siviglia, 14-17 dicembre 2006), Roma, 1713-24.

Fariselli AC, 2013a. Tipi tombali e pratiche funerarie nella necropoli settentrionale di Tharros, in Ocnus. Quaderni della scuola di Specializzazione in Beni Archeologici 18, 9-21.

Fariselli AC, 2013b. Stato sociale e identità nell'Occidente fenicio e punico - I. Le armi in contesto funerario, Biblioteca di Byrsa. Agorà & Co., Lugano.

Fariselli AC, 2014. Ricerche archeologiche e strategie di conservazione nella "necropoli meridionale" di Tharros – Capo San Marco: lo scavo del 2012, in Fariselli (ed.), Da Tharros a Bitia. Nuove prospettive della ricerca archeologica nella Sardegna fenicia e punica. Atti della Giornata di Studio (Bologna 25 marzo 2013) (= DiSCi Archeologia, 3), Bologna, 19-30.

Fariselli AC, 2017. Dinamiche di popolamento a Tharros in età punica. La tomba A2 della necropoli meridionale di Capo San Marco: il contesto archeologico. Byrsa. Scritti sull'antico Oriente mediterraneo, 39-80.

Fariselli AC, Vandini, M., Caillaud, F., Zambruno, S., Caputo, C., Morigi, M.P., Bettuzzi, M., Brancaccio, R., Peccenini E, 2012. Una laminetta in argento con volto virile dalla necropoli meridionale di Tharros. Note iconografiche, archeometriche e conservative. Byrsa Archeol. Punica E Gli Dèi Degli Altri 21-22 23-24, 11–28.

Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, Cherni L, Harich N, Cerny V, Soares P, et al. 2012. The Arabian Cradle: Mitochondrial Relicts of the First Steps along the Southern Route out of Africa. Am. J. Hum. Genet. 90:347–355.

Floris S, 2016. Architettura templare a Tharros – II. Il "Tempio a pianta di tipo semitico" e il "Tempio di Demetra", in Ocnus, n. 24, 47-64.

Floris S. 2014-2015. Architettura templare a Tharros – I. Il "Tempio monumentale" o "delle semicolonne doriche" fra tarda punicità e romanizzazione, Byrsa. Scritti sull'antico Oriente mediterraneo, 39-80.

Fortea J, de la Rasilla M, García-Tabernero A, Gigli E, Rosas A, Lalueza-Fox C. 2008. Excavation protocol of bone remains for Neandertal DNA analysis in El Sidrón Cave (Asturias, Spain). J. Hum. Evol. 55:353–357.

Francalacci P, Morelli L, Angius A, Berutti R, Reinier F, Atzeni R, Pilu R, Busonero F, Maschio A, Zara I, et al. 2013. Low-Pass DNA Sequencing of 1200 Sardinians Reconstructs European Y-Chromosome Phylogeny. Science (80-.). 341:565–569.

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Pääbo S. 2013. DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl. Acad. Sci. U. S. A. 110:2223–7.

Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, et al. 2016. The genetic history of Ice Age Europe. Nature 534:200.

Fulton TL. 2012. Setting Up an Ancient DNA Laboratory. In: Methods in molecular biology (Clifton, N.J.). Vol. 840. p. 1–11.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. 5:5257.

Giglio R, Canzonieri E, 2009. Nuovi dati dalle necropoli ellenistiche e tardo antiche di Lilibeo, in Ampolo C (ed.), Immagine e immagini della Sicilia e di altre isole del Mediterraneo antico. Atti delle seste giornate internazionali di studi sull'area elima e la Sicilia occidentale nel contesto occidentale, vol. II, Pisa, 573-580.

Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I. 2005. Assessing ancient DNA studies. Trends Ecol. Evol. 20:541–4.

Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics 27:2153–2155.

González AM, Karadsheh N, Maca-Meyer N, Flores C, Cabrera VM, Larruga JM. 2008. Mitochondrial DNA variation in Jordanians and their genetic relationship to other Middle East populations. Ann. Hum. Biol. 35:212–231.

Haber M, Doumet-Serhal C, Scheib C, Xue Y, Danecek P, Mezzavilla M, Youhanna S, Martiniano R, Prado-Martinez J, Szpak M, et al. 2017. Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. Am. J. Hum. Genet. 101:274–282.

Haber M, Platt DE, Badro DA, Xue Y, El-Sibai M, Bonab MA, Youhanna SC, Saade S, Soria-Hernanz DF, Royyuru A, et al. 2011. Influences of history, geography, and religion on genetic structure: the Maronites in Lebanon. Eur. J. Hum. Genet. 19:334–40.

Hansen HB, Damgaard PB, Margaryan A, Stenderup J, Lynnerup N, Willerslev E, Allentoft ME. 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. Caramelli D, editor. PLoS One 12:e0170940.

Harvey VL, Egerton VM, Chamberlain AT, Manning PL, Buckley M. 2016. Collagen Fingerprinting: A New Screening Technique for Radiocarbon Dating Ancient Bone. Halcrow SE, editor. PLoS One 11:e0150650.

Hedges REM, Tiemei C, Housley RA. 1992. Results and Methods in the Radiocarbon Dating of Pottery. Radiocarbon 34:906–915.

Hernández CL, Reales G, Dugoujon J-M, Novelletto A, Rodríguez JN, Cuesta P, Calderón R. 2014. Human maternal heritage in Andalusia (Spain): its composition reveals high internal complexity and distinctive influences of mtDNA haplogroups U6 and L in the western and eastern side of region. BMC Genet. 15:11.

Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, et al. 2002. Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups. Am. J. Hum. Genet. 70:1152–1171.

Hervella M, Rotea M, Izagirre N, Constantinescu M, Alonso S, Ioana M, Lazăr C, Ridiche F, Soficaru AD, Netea MG, et al. 2015. Ancient DNA from South-East Europe Reveals Different Events during Early and Middle Neolithic Influencing the European Genetic Heritage. Pereira LMSM, editor. PLoS One 10:e0128810.

Irwin J, Saunier J, Strouss K, Paintner C, Diegoli T, Sturk K, Kovatsi L, Brandstätter A, Cariolou MA, Parson W, et al. 2008. Mitochondrial control region sequences from northern Greece and Greek Cypriots. Int. J. Legal Med. 122:87–89.

Jolliffe IT. 2002. Principal component analysis. Springer.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29:1682–1684.

Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, et al. 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat. Commun. 3:698.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–20.

Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F. 2011. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum. Mutat. 32:25–32.

Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith EA. 2012. Setting the stage – Building and working in an ancient DNA laboratory. Ann. Anat. - Anat. Anzeiger 194:3–6.

Knapp M, Lalueza-Fox C, Hofreiter M. 2015. Re-inventing ancient human DNA. Investig. Genet. 6:4.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15:356.

Krings V, 2000. Quelques considérations sur l'eempire de Carthage». À propos de Malchus, in: Actas Del IV Congreso Internacional de Estudios Fenicios y Punicos (Càdiz, 2 Al 6 de Octubre de 1995). Cadiz, pp. 167–172.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33:1870–1874.

Lazaridis I, Mittnik A, Patterson N, Mallick S, Rohland N, Pfrengle S, Furtwängler A, Peltzer A, Posth C, Vasilakis A, et al. 2017. Genetic origins of the Minoans and Mycenaeans. Nature 548:214.

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. Nature 536:419–424.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–9.

Lindahl T. 1993. Instability and decay of the primary structure of DNA. Nature 362:709–715.

Lindgreen S. 2012. AdapterRemoval: Easy Cleaning of Next Generation Sequencing Reads. BMC Res. Notes 5:337.

Lipinski E, 1995. The Phoenicians, in CANE congress, vol. 2,1321-1333.

Llamas B, Valverde G, Fehren-Schmitz L, Weyrich LS, Cooper A, Haak W. 2017. From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. STAR Sci. Technol. Archaeol. Res. 3:1–14.

Longin R, 1971. New Method of Collagen Extraction for Radiocarbon Dating. Nature 230:241–242.

Loogväli E-L, Roostalu U, Malyarchuk BA, Derenko M V, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk H-V, Parik J, et al. 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. Mol. Biol. Evol. 21:2012–21.

Lorkiewicz W, Płoszaj T, Jędrychowska-Dańska K, Żądzińska E, Strapagiel D, Haduch E, Szczepanek A, Grygiel R, Witas HW. 2015. Between the Baltic and Danubian Worlds: the genetic affinities of a Middle Neolithic population from central Poland. PLoS One 10:e0118316.

Marano M, 2014. L'abitato punico-romano di Tharros (Cabras-OR): i dati d'archivio, in: Fariselli, A.C. (Ed.), Da Tharros a Bitia. Nuove Prospettive Della Ricerca Archeologica Nella Sardegna Fenicia e Punica. Atti Della Giornata Di Studio (Bologna, 25 Marzo 2013). Bononia University Press, Bologna, 75–94

Marano M, 2017. Per una valorizzazione dell'abitato punico-romano di Tharros (Cabras-OR): studio urbanistico preliminare, in: Guirguis, M. (Ed.), From the Mediterranean to the Atlantic: People, Goods and Ideas between East and West. Proceedings of the 8th International Congress of Phoenician and Punic Studies, (Carbonia-Sant'Antioco, 21-26 October 2013). Fabrizio Serra - Editore, Pisa - Roma, 172–177.

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528:499–503.

Mazza F, 1995. Civiltà fenicia e fonti classiche: temi, problemi, prospettive, in I Fenici: Ieri Oggi Domani. Ricerche, scoperte, progetti (Roma 3-5 marzo 1994). Istituto per la Civiltà Fenicia e Punica, Consiglio Nazionale delle Ricerche, Roma.

Melchior L, Kivisild T, Lynnerup N, Dissing J. 2008. Evidence of Authentic DNA from Danish Viking Age Skeletons Untouched by Humans for 1,000 Years. Ahmed N, editor. PLoS One 3:e2214.

Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 2010:pdb.prot5448.

Modi A, Tassi F, Susca RR, Vai S, Rizzi E, Bellis G De, Lugliè C, Gonzalez Fortes G, Lari M, Barbujani G, et al. 2017. Complete mitochondrial sequences from Mesolithic Sardinia. Sci. Rep. 7:42869.

Moscati, S., 1988. Fenicio Punico o Cartaginese. Riv. Studi Fenici 16, 3–13.

Olivieri A, Pala M, Gandini F, Kashani BH, Perego UA, Woodward SR, Grugni V, Battaglia V, Semino O, Achilli A, et al. 2013. Mitogenomes from Two Uncommon Haplogroups Mark Late Glacial/Postglacial Expansions from the Near East and Neolithic Dispersals within Europe. Pereira LMSM, editor. PLoS One 8:e70492.

Olivieri A, Sidore C, Achilli A, Angius A, Posth C, Furtwängler A, Brandini S, Capodiferro MR, Gandini F, Zoledziewska M, et al. 2017. Mitogenome Diversity in Sardinians: A Genetic Window onto an Island's Past. Mol. Biol. Evol. 34:1230–1239.

Ottoni C, Martínez-Labarga C, Loogväli E-L, Pennarun E, Achilli A, De Angelis F, Trucchi E, Contini I, Biondi G, Rickards O. 2009. First Genetic Insight into Libyan Tuaregs: A Maternal Perspective. Ann. Hum. Genet. 73:438–448.

Ottoni C, Ricaut F-X, Vanderheyden N, Brucato N, Waelkens M, Decorte R. 2011. Mitochondrial analysis of a Byzantine population reveals the differential impact of multiple historical events in South Anatolia. Eur. J. Hum. Genet. 19:571–6.

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. 2004. Genetic Analyses from Ancient DNA. Annu. Rev. Genet. 38:645–679.

Pala M, Achilli A, Olivieri A, Hooshiar Kashani B, Perego UA, Sanna D, Metspalu E, Tambets K, Tamm E, Accetturo M, et al. 2009. Mitochondrial haplogroup U5b3: a distant echo of the epipaleolithic in Italy and the legacy of the early Sardinians. Am. J. Hum. Genet. 84:814–21.

Pascucci V, De Falco G, Del Vais C, Sanna I, Melis RT, Andreucci S. 2018. Climate changes and human impact on the Mistras coastal barrier system (W Sardinia, Italy). Mar. Geol. 395:271–284.

Pesce, G., 1966. Tharros. Editrice Sarda F.lli Fossataro, Cagliari.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics 192:1065–93.

Patterson N, Price AL, Reich D. 2006. Population Structure and Eigenanalysis. PLoS Genet. 2:e190.

Pereira L, Prata MJ, Amorim A. 2000. Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. Ann. Hum. Genet. 64:491–506.

Pereira L, Richards M, Goios A, Alonso A, Albarrán C, Garcia O, Behar DM, Gölge M, Hatina J, Al-Gazali L, et al. 2005. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome Res. 15:19–24.

Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, Gerritsen F, Moiseyev V, Gromov A, Raczky P, et al. 2015. Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLoS One 10:e0129102.

Prüfer K, Stenzel U, Hofreiter M, Pääbo S, Kelso J, Green RE. 2010. Computational challenges in the analysis of ancient DNA. Genome Biol. 11:R47.

Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, Morel N, Fernandez-Jalvo Y, Grange T, Geigl E-M. 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. Proc. Natl. Acad. Sci. U. S. A. 104:739–44.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81:559–575.

Quattrocchi Pisano G 1974. I gioielli fenici di Tarros nel Museo Nazionale di Cagliari (= Collezione di Studi Fenici, 3), Roma.

Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford Jr TW, Orlando L, Metspalu E, et al. 2013. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature 505:87–91.

Reimer PJ, Bard E, Bayliss A, Beck JW, Blackwell PG, Ramsey CB, Buck CE, Cheng H, Edwards RL, Friedrich M, et al. 2013. IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon 55:1869–1887.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. Am. J. Hum. Genet. 67:1251–76.

Ringnér M. 2008. What is principal component analysis? Nat. Biotechnol. 26:303–304.

Roostalu U, Kutuev I, Loogväli E-L, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. 2007. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Mol. Biol. Evol. 24:436–48.

Sampietro ML, Gilbert MTP, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C. 2006. Tracking down Human Contamination in Ancient Human Teeth. Mol. Biol. Evol. 23:1801–1807.

Sanna D, Pala M, Cossu P, Dedola GL, Melis S, Fresu G, Morelli L, Obinu D, Tonolo G, Secchi G, et al. 2011. Mendelian breeding units versus standard sampling strategies: Mitochondrial DNA variation in southwest Sardinia. Genet. Mol. Biol. 34:187–94.

Sardo J, 2012. La necropoli punica di Lilibeo (Marsala): studio dei resti osteologici rinvenuti in Corso Gramsci. (A dissertation thesis, Master of Arts. University of Bologna).

Schubert M, Ginolhac A, Lindgreen S, Thompson JF, AL-Rasheid KA, Willerslev E, Krogh A, Orlando L. 2012. Improving ancient DNA read mapping against modern reference genomes. BMC Genomics 13:178.

Schuenemann VJ, Peltzer A, Welte B, van Pelt WP, Molak M, Wang C-C, Furtwängler A, Urban C, Reiter E, Nieselt K, et al. 2017. Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat. Commun. 8:15694.

Secci R, 2016. Nuovi tipi tombali nella necropoli meridionale di Tharros (campagna di scavo 2015), in Byrsa. Scritti sull'antico Oriente mediterraneo 25-26, 27-28 (2014-2015), 185-202.

Sidore C, Busonero F, Maschio A, Porcu E, Naitza S, Zoledziewska M, Mulas A, Pistis G, Steri M, Danjou F, et al. 2015. Genome sequencing elucidates Sardinian genetic architecture and augments association analyses for lipid and blood inflammatory markers. Nat. Genet. 47:1272–1281.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–62.

Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt H-J, Torroni A, Richards MB. 2010. The Archaeogenetics of Europe. Curr. Biol. 20:R174–R183.

Spano G, 1851. Notizie sull'antica città di Tharros, Cagliari, Tipografia Nazionale.

Spanu, P.G., 1998. La Sardegna bizantina tra VI e VII secolo. Mediterraneo Tardoantico e Medievale (= Scavi e Ricerche, 12) Oristano.

Stiller M, Knapp M, Stenzel U, Hofreiter M, Meyer M. 2009. Direct multiplex sequencing (DMPS)--a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. Genome Res. 19:1843–1848.

Terreros MC, Rowold DJ, Mirabal S, Herrera RJ. 2011. Mitochondrial DNA and Y-chromosomal stratification in Iran: relationship between Iran and the Arabian Peninsula. J. Hum. Genet. 56:235–246.

van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. 30:E386–E394.

Usai A, 2014. Alle origini del fenomeno di Mont'e Prama. La civiltà nuragica nel Sinis. In: Minoja, M., Usai, A. (Eds.), Le sculture di Mont'e Prama. Contesto, scavi e materiali, 29–72.

Vella N, 1998. Ritual, landscape and territory: Phoenician and Punic non-funerary religious sites in the Mediterranean; an analysis of the archaeological evidence. (A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Ph.D. in the Faculty of Arts).

Vernesi C, Di Benedetto G, Caramelli D, Secchieri E, Simoni L, Katti E, Malaspina P, Novelletto A, Marin VTW, Barbujani G. 2001. Genetic characterization of the body attributed to the evangelist Luke. Proc. Natl. Acad. Sci. 98:13460–13463.

Vigilant L, Penningtoni R, Harpendingt H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Evolution (N. Y). 86:9350–9354.

Willerslev E, Cooper A. 2005. Ancient DNA. Proc. Biol. Sci. 272:3–16.

Xella P, 2014. «Origini» e «identità», Mélanges de l'École française de Rome - Antiquité, 126–2.

Zucca R, 1984. Tharros. Edizioni Giovanni Corrias, Oristano.

Map of the pre-Roman peoples of Early Italy

CHAPTER 6

Concluding remarks

Technological innovations in the aDNA field, such as the improvement of sequencing methodologies and the analysis of the genomes of ancient peoples, have facilitated the determination of the genealogical relationships between humans as well as the elucidation of migration routes, diversification events and genetic admixture among various groups.

So far, we know that migrations, starting from different areas, have been numerous during the past and have shaped modern models of variation, combined with remote isolation processes: the European prehistory has proved more complex than that which would have been concluded by the models based on the data of the modern-genomic population. For example, due to its strategic geographic position, Italy has long represented a major Mediterranean crossroad where different peoples and cultures mixed over time. However, its multi-layered history of migration pathways and cultural exchanges, has made the reconstruction of its genetic history and population structure extremely controversial and widely debated.

Numerous population movements occurred between the Mediterranean basin and the Middle East during the Metals Ages, a period that have determined the transformation of the first social organizations in ancient civilizations. Nowadays, no sufficient information is available about the origin and possible events of admixture of these populations, and our knowledge is still almost incomplete from a genetic point of view. To date, only few

researches (based on classical low-resolution approach constituted by cloning and Sanger sequencing of the HVS-I mtDNA) have directly investigated the genetic variability in ancient human groups of these periods: the Etruscan and the Nuragic populations. Briefly, the analyses of the mtDNA (HVS-I) in the Etruscan human remains (Vernesi et al., 2004; Ghirotto et al., 2013), showed the presence of genetic characteristics typical of the Near East, while the comparison with the modern populations has revealed some evolutionary continuity between the Etruscans and the present-day inhabitants of Casentino and Volterra in the Tuscan region (Ghirotto S et al., 2013). However, the computational analyses and the demographic simulations carried out place the link between Tuscany and Anatolia at a remote stage of prehistory, which could be traced to the spread of farmers during the Neolithic period (~6500 ya) (Ghirotto et al., 2013; Rates et al., 2013), thus leaning for a local origin of the Etruscans from probably the previous 'Villanovan culture'. About the Sardinian population, 23 Bronze Age and Iron Age Nuragic individuals (2,700-3,430 BC), collected from different sites across Sardinia island, were characterised by low mtDNA diversity (HVS-I) and by a lack of geographical and temporal genetic structure when compared to current Central Sardinians of the Ogliastra province (Caramelli et al., 2007). The analysis showed a direct genealogical continuity between Sardinians and the modern people of Ogliastra, but not Gallura, which has a much higher probability than any alternative scenarios and that genetic diversity in Gallura evolved largely independently, owing in part to gene flow from the mainland (Ghirotto et al., 2010).

In this work, I focused the attention on the Italian Iron Age populations of the Piceni from Novilara necropolis, Marche region (Case Study I) and on the Punic population from the 'southern necropolis' of Tharros (OR) located in Sardinia and from the necropolis of Lilybaeum (TP) in Sicily (Case Study II), never studied before from a genetic point of view.

In the Case Study I, the apparent matrilineal genetic continuity between the ancient Piceni and modern populations of the Marche region (Ancona, Macerata, and Ascoli Piceno) suggests that probably the different migratory events that involved this area during historical period did not influence the maternal gene pool of their inhabitants. As suggest by Günther and Jakobsson (2016) after the turnovers during the Early Neolithic and Bronze Age periods, the genetic composition of populations in some areas of Europe were starting to become alike to the present-day people of the same regions. This consideration does not negate successive migrations, but suggests that the populations involved were not as highly varied as during the Neolithic (Lazaridis et al., 2016).

Probably, in Central Italy, there was not such a strong reshuffling in the maternal genetic pool during historical periods, when, for instance, migrations due to Celts, Romans and Goths are attested. Indeed, as suggested by previous studies on present-day Italian population, the actual sex-biased genetic structure in Italy is possibly the result of different demographic histories for males and females, with the more homogenous pattern of mtDNA variability probably tracing back to more ancient times, and the Y-chromosome structure being instead shaped by more recent migration events (Boattini et al., 2013).

In the Case Study II, the analysis of the HVS-I of the mtDNA of Tharros individuals (methodological approach 1), showed that they are genetically closer to the current inhabitants of the North Africa area (Tunisia), rather the modern-day Sardinian populations. Although preliminary and being aware that the matrilineal DNA perspective may only reveal a part of the population history, the results seems to indicate in the Tharros population, at least with regard to the historical period of Carthaginian control in Sardinia, a greater incidence of the North African component than the autochthonous one. This result is in accordance with the archaeological evidence coming from one of the two graves (grave A2), whose grave goods (ceramics fragment) and the morphology of the tombs was recently published by Fariselli (2017). It was possible to obtain one whole genome (methodological approach 2) from a Tharros sample dated back to the 382 ± 153 BC. By comparing that genome with ancient reference populations, individuals from the European/Anatolian Early Neolithic and European Middle Neolithic/Chalcolithic appear to be those genetically closer. When restricting to present-day reference individuals, Tharros sample (DA393) seems to be genetically related to the current Sardinians, rather to the modern Phoenician-influenced populations of the Mediterranean basin, such as Middle/Near East, North Africa, Sicily, Spain, and Cyprus, suggests that DA393 sample probably was a 'Sardinian autochthonous'. The closeness between the Punic sample (4th-3th century BC) and the present-day Sardinians suggests a certain degree of genetic continuity, probably due to the isolated position of the island: the contemporary Sardinians harbour a unique genetic heritage, as a result of their distinct history and relative isolation from the demographic upheavals of continental Europe. In a new study based on the complete mitochondrial genome, the researchers have tried to clarify the origins of the Sardinian population in the context of European prehistory and ancient human migrations (Olivieri et al., 2017). The results suggested that 78.4 % of the modern-day Sardinian mitogenomes belong to branches that cannot be found anywhere else outside the island. Thus, they were defined as Sardinian-Specific Haplogroups (SSHs) that

most likely arose in the island after its initial occupation. Most of them appear to have descended from the first farmers who occupied the island since the Neolithic and Copper Age. However, the finding of some rare SSHs do not completely discard the hypothesis that another population already lived on the island prior to the Neolithic, a scenario that would agree with archaeological evidence of a Mesolithic occupation of Sardinia (Olivieri et al 2017).

Although this work contributes to enlarge the knowledge of the Italian pre-historic populations, in order to better clarify the complex situation highlighted here, it would be necessary to enlarge the set of Italian (but not only) ancient genomes from the same historical periods.

References:

Boattini A, Martinez-Cruz B, Sarno S, Harmant C, Useli A, Sanz P, Yang-Yao D, Manry J, Ciani G, Luiselli D, et al. 2013. Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata. PLoS One 8:e65441.

Caramelli D, Vernesi C, Sanna S, Sampietro L, Lari M, Castrì L, Vona G, Floris R, Francalacci P, Tykot R, et al. 2007. Genetic variation in prehistoric Sardinia. Hum. Genet. 122:327–36.

Fariselli AC, 2017. Dinamiche di popolamento a Tharros in età punica. La tomba A2 della necropoli meridionale di Capo San Marco: il contesto archeologico. Byrsa. Scritti sull'antico Oriente mediterraneo, 39-80.

Ghirotto S, Mona S, Benazzo A, Paparazzo F, Caramelli D, Barbujani G. 2010. Inferring genealogical processes from patterns of Bronze-Age and modern DNA variation in Sardinia. Mol. Biol. Evol. 27:875–86.

Ghirotto S, Tassi F, Fumagalli E, Colonna V, Sandionigi A, Lari M, Vai S, Petiti E, Corti G, Rizzi E, et al. 2013. Origins and Evolution of the Etruscans' mtDNA. Hawks J, editor. PLoS One 8:e55519.

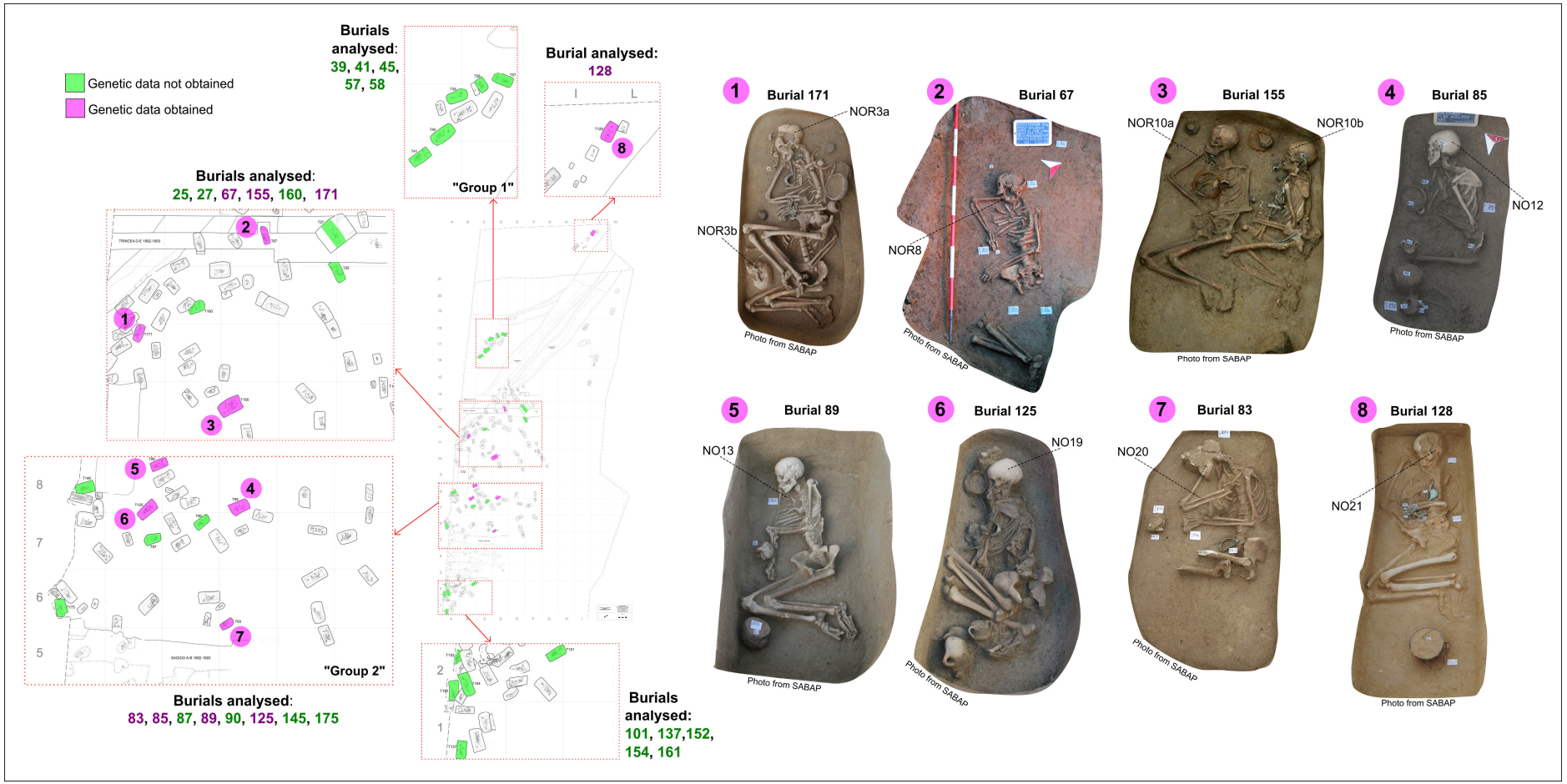
Günther T, Jakobsson M. 2016. Genes mirror migrations and cultures in prehistoric Europe — a population genomic perspective. Curr. Opin. Genet. Dev. 41:115–123.

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. Nature 536:419–424.

Olivieri A, Sidore C, Achilli A, Angius A, Posth C, Furtwängler A, Brandini S, Capodiferro MR, Gandini F, Zoledziewska M, et al. 2017. Mitogenome Diversity in Sardinians: A Genetic Window onto an Island's Past. Mol. Biol. Evol. 34:1230–1239.

Vernesi C, Caramelli D, Dupanloup I, Bertorelle G, Lari M, Cappellini E, Moggi-Cecchi J, Chiarelli B, Castrì L, Casoli A, et al. 2004. The Etruscans: a population-genetic study. Am. J. Hum. Genet. 74:694–704.

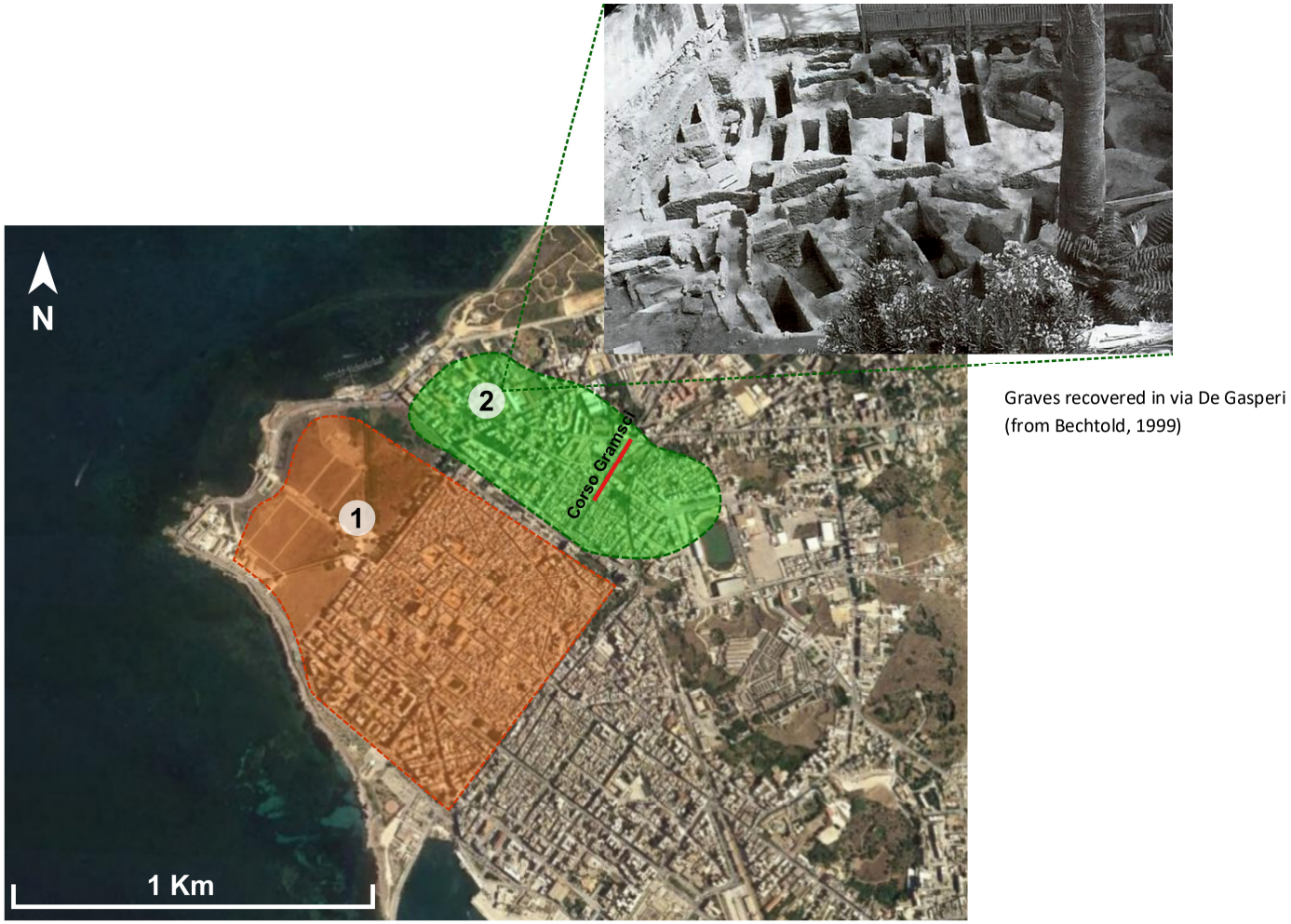
Appendix I



S-Figure 4.4.4.1.2 | Map of Novilara archaeological site. The burials of the individuals selected for the molecular analyses are highlighted in green colour. Photos of the individuals whose genetic profiles were obtained (in pink colour) are reported on the right side. Abbreviations: SABAP, Archivio Soprintendenza Archeologia Belle Arti e Paesaggio delle Marche.

Appendix II

S-Figure II.1| Archaeological site of Tharros. The skeletal remains analysed in this study were collected from the graves T.2 (2012); T.Z and T.L (2013); T.A2 (2014); T.B7, T. B10 and T. B14 (2015); T.B14 (2016).



S-Figure II.2| The area of Lilibaeum: urban site (1) and the necropolis (2). Samples analysed in this study come from the graves found in Corso Gramsci road (red line).

S-Table 5.2.5.7.1.1a | List of the 1,753 individuals from 82 Euro-Mediterranean populations included in the modern reference dataset used for the genome-wide SNP analyses.

Population	*N*	*Reference*
Mozabite	23	Li et al. 2008
Algerian_Jews	5	Behar et al. 2012
Moroccans	8	Behar et al. 2010
Moroccan_Jews	18	Behar et al. 2010; Behar et al. 2012
Libyan_Jews	6	Behar et al. 2012
Tunisian_Jews	5	Behar et al. 2012
Egyptans	11	Behar et al. 2010
Saudis	15	Behar et al. 2010
Yemenese	6	Behar et al. 2010
Yemenite_Jews	18	Behar et al. 2010; Behar et al. 2012
Jordanians	18	Behar et al. 2010
Syrians	15	Behar et al. 2010
Syrian_Jews	2	Behar et al. 2012
Lebanese	8	Behar et al. 2010
Samaritians	2	Behar et al. 2010
Palestinian	46	Li et al. 2008; Behar et al. 2012
Bedouin	44	Li et al. 2008
Druze	41	Li et al. 2008; Behar et al. 2012
Turks	18	Behar et al. 2010
Iranians	15	Behar et al. 2010
Iranian_Jews	12	Behar et al. 2010; Behar et al. 2012
Iraqi_Jews	13	Behar et al. 2010; Behar et al. 2012
Azeris	22	Yunusbayev et al. 2013
Armenians	31	Behar et al. 2010; Yunusbayev et al. 2011
Georgians	27	Behar et al. 2010; Behar et al. 2012
Georgian_Jews	6	Behar et al. 2010; Behar et al. 2012
Kurdish_Jews	9	Behar et al. 2012
Azerbaijani_Jews	11	Behar et al. 2010; Behar et al. 2012
Abhkasians	22	Yunusbayev et al. 2011; Behar et al. 2012
Balkars	20	Yunusbayev et al. 2011; Yunusbayev et al. 2013
Chechens	19	Yunusbayev et al. 2011
North_Ossetians	18	Yunusbayev et al. 2011; Behar et al. 2012
Tabasaran	3	Behar et al. 2012
Nogais	16	Yunusbayev et al. 2011
Kumyks	16	Yunusbayev et al. 2011
Kabardins	3	Yunusbayev et al. 2013
Lezgins	21	Behar et al. 2010; Behar et al. 2012
Adygei	15	Li et al. 2008
Sephardic_Jews	21	Behar et al. 2010; Behar et al. 2012
Ashkenazy_Jews	24	Behar et al. 2010; Behar et al. 2012
Italian_Jews	10	Behar et al. 2012

Sardinian	28	Li et al. 2008
Italian_Bergamo	13	Li et al. 2008
Italian_Tuscan	8	Li et al. 2008
TSI	107	1000 Genomes
Italian_Abruzzo	11	Behar et al. 2012
Italian_Sicily	13	Behar et al. 2012
Cypriots	12	Behar et al. 2012
Greeks	32	Behar et al. 2012; Kushniarevich et al. 2015
Macedonians	14	Kovacevic et al. 2014
Bulgarians	13	Yunusbayev et al. 2011
Romanians	14	Behar et al 2010
Kosovars	9	Kovacevic et al. 2014
Montenegrins	14	Kovacevic et al. 2014
Croats	23	Behar et al. 2012
Bosnians	14	Kovacevic et al. 2014
Serbians	18	Kovacevic et al. 2014
Slovenians	15	Kushniarevich et al. 2015
Hungarians	19	Behar et al. 2010
Slovaks	15	Kushniarevich et al. 2015
Ukranians	20	Yunusbayev et al. 2011
Moldavians	7	Behar et al. 2012
Belorussians	17	Behar et al. 2010; Behar et al. 2012; Kushniarevich et al. 2015
Russians	65	Li et al. 2008; Behar et al. 2010; Behar et al. 2012; Yunusbayev et al. 2013; Kushniarevich et al. 2015
Mordovians	15	Yunusbayev et al. 2011
Gagauzes	12	Yunusbayev et al. 2013
Vepsas	11	Yunusbayev et al. 2013
Karelians	14	Yunusbayev et al. 2013
Estonians	21	Raghavan et al. 2014; Kushniarevich et al. 2015
Latvians	6	Kushniarevich et al. 2015
Lithuanians	9	Behar et al. 2010
Polish	17	Behar et al. 2012
Germans	13	Yunusbayev et al. 2013
French	27	Li et al. 2008
French_Basque	24	Li et al. 2008
Spaniards	12	Behar et al. 2010
IBS	95	1000 Genomes
CEU	96	1000 Genomes
GBR	89	1000 Genomes
Orcadian	13	Li et al. 2008
Swedish	18	Behar et al. 2012
FIN	97	1000 Genomes

S-Table 5.2.5.7.1.1b | List of the 390 ancient samples extracted from the literature included in the comparisons.

ID	*Ancient reference group*
Ust_Ishim	Ust_Ishim
I0876	Kostenki14
I0898	Kostenki12
I0909.damage	Muierii2
Cioclovina_d	Cioclovina1
MA1	MA1
I9050.damage	AfontovaGora3
AG2	AfontovaGora2
KK1	CHG
SATP	CHG
I0004.damage	Vestonice_Cluster
I0080.damage	Vestonice_Cluster
I0062	Vestonice_Cluster
I0065.damage	Vestonice_Cluster
I0066.damage	Vestonice_Cluster
I0869.damage	Vestonice_Cluster
I1577	Vestonice_Cluster
GA252snp	Vestonice_Cluster
I0006_damage	Vestonice14
I0889.damage	Ostuni2
B1_d	Paglicci108
Q376-19_d	GoyetQ376-19
Q56-16_d	GoyetQ56-16
Q53-1_d	GoyetQ53-1
Q116-1	GoyetQ116-1
Q2	ElMiron_Cluster
I0907.damage	ElMiron_Cluster
HF49	ElMiron_Cluster
Hohle_Fels	ElMiron_Cluster
Rigney2	ElMiron_Cluster
BUR_d	ElMiron_Cluster
BRI_d	Brillenhohle
Ofnet.damage	Ofnet
Bockstein.damage	Bockstein
LCX-13.damage	Villabruna_Cluster
Falkenstein_d	Villabruna_Cluster
Berry_au_Bac	Villabruna_Cluster
CRC-1_d	Villabruna_Cluster
Rochedane	Villabruna_Cluster
Ranchot.damage	Villabruna_Cluster
I9030	Villabruna_Cluster

ADI_d	Iboussieres39
I0878.damage	Continenza
LaBrana1	WHG
I1507	WHG
Loschbour	WHG
Bichon	Bichon
I0013	SHG
I0011	SHG
I0015	SHG
I0012	SHG
I0014	SHG
Motala12	SHG
I0124	EHG
I0211	EHG
I0061	EHG
I0434	Samara_Eneolithic
I0433	Samara_Eneolithic
I0122	Samara_Eneolithic
RISE240	Yamnaya_Kalmykia
RISE546	Yamnaya_Kalmykia
RISE547	Yamnaya_Kalmykia
RISE548	Yamnaya_Kalmykia
RISE550	Yamnaya_Kalmykia
RISE552	Yamnaya_Kalmykia
I0231	Yamnaya_Samara
I0370	Yamnaya_Samara
I0441	Yamnaya_Samara
I0444	Yamnaya_Samara
I0439	Yamnaya_Samara
I0357	Yamnaya_Samara
I0429	Yamnaya_Samara
I0438	Yamnaya_Samara
I0443	Yamnaya_Samara
RISE507	Afanasievo
RISE508	Afanasievo
RISE509	Afanasievo
RISE510	Afanasievo
RISE511	Afanasievo
I0371	Poltavka
I0126	Poltavka
I0440	Poltavka
I0374	Poltavka
I0418	Potapovka
I0419	Potapovka
I0246	Potapovka

RISE555	Russia_EBA
I0354	Srubnaya_outlier
I0360	Srubnaya_outlier
I0423	Srubnaya_outlie
I0432	Poltavka_outlier
I0235	Srubnaya
I0234	Srubnaya
I0431	Srubnaya
I0430	Srubnaya
I0424	Srubnaya
I0232	Srubnaya
I0358	Srubnaya
I0361	Srubnaya
I0359	Srubnaya
I0422	Srubnaya
RISE386	Sintashta
RISE391	Sintashta
RISE392	Sintashta
RISE394	Sintashta
RISE395	Sintashta
RISE500	Andronovo
RISE503	Andronovo
RISE505	Andronovo
RISE512	Andronovo_outlier
I0247	Scythian_IA
I0156.SG	England_ancient
I0157.SG	England_ancient
I0159.SG	England_ancient
I0160.SG	England_ancient
I0161.SG	England_ancient
I0769.SG	England_ancient
I0773.SG	England_ancient
I0774.SG	England_ancient
I0777.SG	England_ancient
I0789.SG	England_ancient
3DT16.SG	England_ancient
6DT18.SG	England_ancient
6DT21.SG	England_ancient
6DT22.SG	England_ancient
6DT23.SG	England_ancient
6DT3.SG	England_ancient
M1489.SG	England_ancient
NO3423.SG	England_ancient
3DT26.SG	England_Roman_MiddleEast
rath1.SG	Europe_LNBA

rath2.SG	Europe_LNBA
rath3.SG	Europe_LNBA
RISE471	Central_LNBA_outlier
I0047	Central_LNBA
I0099	Central_LNBA
I0171	Central_LNBA
I0059	Central_LNBA
I1542	Central_LNBA
I1536	Central_LNBA
I1544	Central_LNBA
I1538	Central_LNBA
I1539	Central_LNBA
I1534	Central_LNBA
I0106	Central_LNBA
I1540	Central_LNBA
I1532	Central_LNBA
I0049	Central_LNBA
I0550	Central_LNBA
I0115	Central_LNBA
I0117	Central_LNBA
I0804	Central_LNBA
I0803	Central_LNBA
I0164	Central_LNBA
RISE00	Central_LNBA
RISE109	Central_LNBA
RISE150	Central_LNBA
RISE154	Central_LNBA
RISE431	Central_LNBA
RISE434	Central_LNBA
RISE435	Central_LNBA
RISE436	Central_LNBA
RISE446	Central_LNBA
RISE577	Central_LNBA
RISE586	Central_LNBA
I0103	Central_LNBA
I0104	Central_LNBA
I0116	Central_LNBA
I0118	Central_LNBA
I1546	Bell_Beaker_LN
I0806	Bell_Beaker_LN
I0805	Bell_Beaker_LN
I0113	Bell_Beaker_LN
I0112	Bell_Beaker_LN
I0060	Bell_Beaker_LN
I0111	Bell_Beaker_LN

I0108	Bell_Beaker_LN
RISE559	Bell_Beaker_LN
RISE560	Bell_Beaker_LN
RISE562	Bell_Beaker_LN
RISE563	Bell_Beaker_LN
RISE564	Bell_Beaker_LN
RISE566	Bell_Beaker_LN
RISE568	Bell_Beaker_LN
RISE569	Bell_Beaker_LN
I1549	Bell_Beaker_LN
RISE47	Northern_LNBA
RISE61	Northern_LNBA
RISE71	Northern_LNBA
RISE94	Northern_LNBA
RISE97	Northern_LNBA
RISE98	Northern_LNBA
RISE175	Northern_LNBA
RISE179	Northern_LNBA
RISE210	Northern_LNBA
RISE276	Northern_LNBA
I1502	Hungary_BA
I1504	Hungary_BA
RISE247	Hungary_BA
RISE254	Hungary_BA
RISE349	Hungary_BA
RISE371	Hungary_BA
RISE373	Hungary_BA
RISE374	Hungary_BA
RISE479	Hungary_BA
RISE480	Hungary_BA
RISE483	Hungary_BA
RISE484	Hungary_BA
ATP9	Iberia_BA
I1282	Iberia_Chalcolithic
I1276	Iberia_Chalcolithic
I1284	Iberia_Chalcolithic
I1280	Iberia_Chalcolithic
I1314	Iberia_Chalcolithic
I1277	Iberia_Chalcolithic
I1272	Iberia_Chalcolithic
I1281	Iberia_Chalcolithic
I1300	Iberia_Chalcolithic
I1271	Iberia_Chalcolithic
I1303	Iberia_Chalcolithic
ATP2	Iberia_Chalcolithic

ATP16	Iberia_Chalcolithic
Matojo	Iberia_Chalcolithic
bally.SG	Europe_MNChL
I0405	Iberia_MN
I0407	Iberia_MN
I0408	Iberia_MN
I0406	Iberia_MN
I0807	Central_MN
I0559	Central_MN
I0560	Central_MN
I0551	Central_MN
I1497	Central_MN
I0172	Central_MN
RISE486	Remedello
RISE487	Remedello
RISE489	Remedello
Iceman	Iceman
I0409	Iberia_EN
I0412	Iberia_EN
I0410	Iberia_EN
I0413	Iberia_EN
CB13	Iberia_EN
I0046	LBK_EN
I0048	LBK_EN
I0057	LBK_EN
I0100	LBK_EN
I0659	LBK_EN
I0821	LBK_EN
I1550	LBK_EN
I0797	LBK_EN
I0795	LBK_EN
I0022	LBK_EN
I0026	LBK_EN
I0025	LBK_EN
I0054	LBK_EN
Stuttgart	Stuttgart
I0176	LBKT_EN
I0174	Hungary_EN
I1508	Hungary_EN
I1500	Hungary_EN
I1499	Hungary_EN
I1495	Hungary_EN
I1498	Hungary_EN
I1506	Hungary_EN
I1496	Hungary_EN

I1505	Hungary_EN
Bar31.SG	Anatolia_N
Bar8.SG	Anatolia_N
I1581	Anatolia_N
I1583	Anatolia_N
I1580	Anatolia_N
I1585	Anatolia_N
I1579	Anatolia_N
I1100	Anatolia_N
I1102	Anatolia_N
I1099	Anatolia_N
I1103	Anatolia_N
I1101	Anatolia_N
I1097	Anatolia_N
I0744	Anatolia_N
I1096	Anatolia_N
I1098	Anatolia_N
I0708	Anatolia_N
I0745	Anatolia_N
I0746	Anatolia_N
I0707	Anatolia_N
I0709	Anatolia_N
I0736	Anatolia_N
I0726	Anatolia_N
I0723	Anatolia_N
I0724	Anatolia_N
I0727	Anatolia_N
I0725	Anatolia_N_outlier
Bon001	Anatolia_Boncuklu
Bon002	Anatolia_Boncuklu
Bon004	Anatolia_Boncuklu
Bon005	Anatolia_Boncuklu
Tep002	Anatolia_Tepecik_Ciftlik
Tep003	Anatolia_Tepecik_Ciftlik
Tep004	Anatolia_Tepecik_Ciftlik
Tep006	Anatolia_Tepecik_Ciftlik
Tep001	Anatolia_Tepecik_Ciftlik_outlier
kum4.SG	Anatolia_Kumtepe
kum6.SG	Anatolia_Kumtepe
I2683	Anatolia_BA
I2495	Anatolia_BA
I2499	Anatolia_BA
I1584	Anatolia_ChL
I1072	Natufian
I1069	Natufian

I1687	Natufian
I1690	Natufian
I1685	Natufian
I0861	Natufian
I1679	Levant_N
I1416	Levant_N
I1415	Levant_N
I1414	Levant_N
I1701	Levant_N
I1709	Levant_N
I1727	Levant_N
I1710	Levant_N
I1707	Levant_N
I1704	Levant_N
I1700	Levant_N
I1699	Levant_N
I0867	Levant_N
I1705	Levant_BA
I1706	Levant_BA
I1730	Levant_BA
AH1.SG	Iran_N
AH2.SG	Iran_N
AH4.SG	Iran_N
WC1.SG	Iran_N
I1290	Iran_N
I1944	Iran_N
I1945	Iran_N
I1949	Iran_N
I1951	Iran_N
I1293	Iran_HotuIIIb
I1671	Iran_LN
I1661	Iran_ChL
I1670	Iran_ChL
I1662	Iran_ChL
I1674	Iran_ChL
I1665	Iran_ChL
F38.SG	Iran_IA
I1955	Iran_recent
I1635	Armenia_EBA
I1633	Armenia_EBA
I1658	Armenia_EBA
I1656	Armenia_MLBA
RISE396	Armenia_MLBA
RISE397	Armenia_MLBA
RISE407	Armenia_MLBA

RISE408	Armenia_MLBA
RISE412	Armenia_MLBA
RISE413	Armenia_MLBA
RISE416	Armenia_MLBA
RISE423	Armenia_MLBA
I1634	Armenia_ChL
I1632	Armenia_ChL
I1631	Armenia_ChL
I1409	Armenia_ChL
I1407	Armenia_ChL
I0070	Minoan_Lasithi
I0071	Minoan_Lasithi
I0073	Minoan_Lasithi
I0074	Minoan_Lasithi
I9005	Minoan_Lasithi
I9127	Minoan_Odigitria
I9128	Minoan_Odigitria
I9129	Minoan_Odigitria
I9130	Minoan_Odigitria
I9131	Minoan_Odigitria
I9123	Crete_Armenoi
I9010	Mycenaean
I9006	Mycenaean
I9033	Mycenaean
I9041	Mycenaean
Klei10.SG	Greece_N
Pal7.SG	Greece_N
Rev5.SG	Greece_N
I2937	Greece_N
ERS1790729	Sidon_BA
ERS1790730	Sidon_BA
ERS1790731	Sidon_BA
ERS1790732	Sidon_BA
ERS1790733	Sidon_BA